The Dissertation Committee for Gautam Sakleshpur Muralidhar
certifies that this is the approved version of the following dissertation:

Computer-Aided Analysis and Interpretation of Breast Imaging Data

Committee:

Alan C. Bovik, Supervisor

Mia K. Markey, Supervisor

Wilson S. Geisler III

Kristen Grauman

H. Grady Rylander, III

Gary J. Whitman

Computer-Aided Analysis and Interpretation of Breast Imaging Data

by

Gautam Sakleshpur Muralidhar, B.E.; M.S.E.

DISSERTATION

Presented to the Faculty of the Graduate School of
The University of Texas at Austin
in Partial Fulfillment
of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN
December 2012

Dedicated to my wife Sindhu, my father Muralidhar, and my late uncle Sitaram.

Acknowledgments

This amazing journey would not have been possible had it not been for the support and encouragement of several people. Firstly, I would like to express my deepest gratitude to my wife Sindhu who has been through this journey with me. Sindhu was a constant source of strength for me, without which this journey would not have been completed. A Ph.D. degree invariably takes a student through several crests and troughs and I was no different. Sindhu, being a Ph.D. student herself knew this very well. She kept me humble while riding a crest, and used to perk up my spirits when going through a trough. Most importantly, she made me believe in myself. This journey with Sindhu was made extra special with the birth of our first son, Vivan, who at the time of writing this is just 27 days old. He has been a great source of joy for us and we are thankful to the almighty for this wonderful gift.

I would also like to express my sincere gratitude to my family members. My father, Muralidhar, has been a great source of strength and moral support to me during this journey. No words can describe the sacrifices he made while raising me as a child and this Ph.D. is in many ways a fruit of his hard work. My gratitude goes out to my late uncle, Sitaram, and my aunt Janaki for always being there for me. My uncle was a scientist and he was an inspiration to me to pursue this Ph.D. Many thanks to their sons Ravi and Kiran, and

their respective families for all the wonderful times.

I would like to sincerely thank my Ph.D. advisers Dr. Al Bovik and Dr. Mia Markey for the huge role they played in this journey. I consider myself to be extremely fortunate to have worked with Al and Mia. Outside of my family, I can't think of any other two people who have supported and encouraged me like the way Al and Mia have done. There were times during my Ph.D. when I thought I would be caught in a trough forever. The unflinching support and advice of Al and Mia were crucial in helping me get out of such situations. I will always be grateful to their backing while working on some challenging problems during the course of my Ph.D. Special thanks to Al for all the wonderful, thought provoking discussions on stereo and the singularity index at Starbucks. These discussions helped me immensely in developing my ideas. Further, the passion and commitment shown by Al and Mia towards science and education is absolutely inspirational. Every second I spent with them was a wonderful learning experience that I will cherish for the rest of my life. I would also like to thank them both for giving me the opportunity to work on several problems outside of my dissertation, which helped me to develop exposure to other interesting applications of image processing and computer vision. Finally, I would like to thank Mia and her wonderful family for all their help and advice on baby matters.

During the course of my Ph.D., I was privileged to work with Dr. Gary Whitman at The University of Texas MD Anderson Cancer Center (UT MDACC). I regard Gary as my clinical mentor and I have learned a lot about breast cancer and breast imaging from him. I am extremely grateful to Gary for always being very proactive to my research needs. His enthusiasm towards my research is something that I will fondly remember for the rest of my life. I am also grateful to Gary for taking time of his busy schedule and driving all the way down from Houston to Austin to be here physically for both my proposal exam as well as my defense.

I would like to thank my committee members Dr. Bill Geisler, Dr. Kristen Grauman, and Dr. Grady Rylander for their valuable suggestions on my dissertation. Apart from knowing them as my committee members, I have also had the privilege of taking their classes. Their teaching and knowledge of their subject has been inspirational and no words are enough to describe the vast amount of knowledge I have gained from their classes.

I would like to thank Dr. Joydeep Ghosh with whom I had the good fortune of working on the problem of image quality assessment. I learned a lot about data mining concepts such as topic modeling from Joydeep. Working with him has been an invaluable experience for me since it gave me the opportunity to broaden my skill set.

I would also like to thank the many radiologists at UT MDACC for their participation in my research projects. Special thanks go out to Dr. Tamara Haygood and Dr. Tanya Stephens. I have worked with both Tamara and Tanya on multiple projects. They have always been responsive to my research questions and have been very helpful with administrative matters such as IRB approval.

I have made many new friends at UT who have made this journey extra special. Many thanks to the past and current members of LIVE, particularly, Mehul, Sumo, Anush, Rajiv, Ajay, Anish, Ming, Tony, Lark, Michele, Dinesh, Deepti, and Greg. I have learned a lot from each of them, and in particular, I have had an amazing time working with Anish and Rajiv on image quality assessment. I would also like to thank the past and current members of BMIL, particularly, Wendy, Daifeng, Shuang, Juhun, Nishant, Clement, Gezheng, Tejas, Nisha, and Ricky. Special thanks to Juhun and Nishant for their collaborations. I will also cherish the fun times I have spent with Mehul and his wife Pallavi, and Juhun and his wife Eunjung for the rest of my life.

I would like to express my gratitude to the administrative staff in Biomedical Engineering and in Electrical and Computer Engineering. I would specially like to thank Heidi, Carol, Krystal, Michael, Margo, and Janet. Finally, I would also like to express my sincere thanks to the UT Library System, without which none of the work in this dissertation would have been possible.

Gautam Sakleshpur Muralidhar

The University of Texas at Austin
December 2012

Computer-Aided Analysis and Interpretation of Breast Imaging Data

Publication No. ____________________

Gautam Sakleshpur Muralidhar, Ph.D.
The University of Texas at Austin, 2012

Supervisors: Alan C. Bovik
Mia K. Markey

Early detection of breast cancer on screening mammograms is crucial to reduce mortality rates. Computer-aided detection (CADe) systems for mammography are of great importance since they have been shown to positively assist radiologists in detecting early cancer. However, one area where CADe systems for mammography need improvement is in the early detection and annotation of spiculated lesions, which may represent invasive malignancies, and hence, early detection is crucial. Spicule annotation is important since it can yield useful discriminative information about the suspect lesion location on the mammogram and can also provide rich visual evidence to the interpreting radiologist to make the right follow-up decision. However, spicule annotation is a non-trivial task since spicules are fine scale curvilinear structures that are often not clearly visible amidst the surrounding breast parenchyma. The first contribution of this dissertation is an active contour algorithm called *snakules*

for the annotation of spicules on mammography. Observer studies with experienced radiologists to evaluate the performance of snakules demonstrate the potential of the algorithm as an annotation tool that could be used to augment existing spiculated mass CADe systems.

Mammography suffers from a major limitation: the 3-D to 2-D projection process results in anatomical noise due to overlapping of out of plane tissue structures, which hinders both radiologists and CADe systems in finding early cancers. This has motivated the development of 3-D breast imaging in the form of breast tomosynthesis, stereoscopic (stereo) mammography, and breast computed tomography (CT) to augment mammography for early cancer detection. Our second contribution is a novel computational stereo model for estimating a dense disparity map from a pair of stereo mammograms. This problem is very important since this is the first step towards elucidating 3-D information that is essential for conducting 3-D digital analysis on the stereo mammogram images. Nearly all of the 3-D structural information of interest on a stereo mammogram exists as a complex network of multi-layered, heavily occluded curvilinear structures, which is unlike what is seen on optical images of the real world. Our proposed stereo model employs a new singularity index as a constraint in a global optimization framework to obtain better estimates of disparity along critical curvilinear structures. The new singularity index is an important contribution of this work. In-depth theoretical analyses and experiments on several real world images demonstrate the efficacy of the index for detecting multi-scale curvilinear structures. Experiments on synthetic

images with known ground truth and on real stereo mammograms highlight the advantages of the proposed stereo model over the canonical stereo model.

The final contribution of this dissertation is an observer study, which demonstrates the feasibility of viewing breast tomosynthesis projection images stereoscopically. Unlike stereo mammogram images, each tomosynthesis projection image is acquired at a much lower dose. Stereo viewing of tomosynthesis projection images has the potential to reveal the 3-D structure of the breast, unlike the current cine or slice-by-slice viewing modes. The results from our study suggest that stereo viewing could be a viable reading mode for breast tomosynthesis data in the future.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Significance

Breast cancer is the second most frequently diagnosed form of cancer in women in United States after skin cancer. The American Cancer Society estimates that there will be 226,870 new cases of breast cancer in women in 2012. Among cancers, breast cancer is also the second leading cause of death in women after lung cancer. It is estimated by the American Cancer Society that 39,510 women will die in 2012 due to breast cancer. Early detection of breast cancer is crucial to reduce mortality rates. Mammography remains the first choice modality for screening asymptomatic women for detecting signs of early breast cancer. Excellent image resolution ($<$ 100 microns per pixel) at a low radiation dose, reasonable cost, short image acquisition time, and ease of use have made mammography practical. However, mammography suffers from one major drawback: the loss of 3-D information due to the projection of the breast onto a 2-D image plane. The 3-D to 2-D projection process results in what is commonly referred to as anatomical noise due to overlapping of out of plane tissue structures. Anatomical noise is one of the key factors hindering the correct interpretation of a mammogram. For example, anatomical noise could obscure subtle cancers, or could cause false visual relationships between different tissue structures suggesting a cancer, where in reality there may exist none. Such inaccurate diagnoses often result in additional imaging tests and biopsy procedures that add to both monetary and emotional costs for women undergoing these procedures. In fact, the positive predictive value (defined as the ratio of the number of confirmed cancers to the number women called back for further diagnostic work-up) of mammography in a routine screening program is quite low and is reported to be around 10-30% [46],[97].

Computer-aided detection (CADe) systems have been developed to assist radiologists in interpreting mammograms [94], [77], and [21]. These systems are used as second readers in a clinical setting and commercial CADe systems have been in use for more than a decade now. A number of clinical [32], [35], [98], as well as a long history of laboratory studies have demonstrated the high potential of CADe for detecting early cancer.

In spite of the promise shown by CADe systems in detecting breast cancer early, there are certain problems that need to be addressed for the systems to realize their full potential. The studies by Fenton *et al.* [27], [28] have raised a few uncomfortable questions on the current state of CADe systems in clinical practice. Fenton *et al.* suggested that there was no concrete evidence supporting the claim that CADe helped in reducing mortality rates in women with invasive breast cancers , even though CADe had shown some promise in detecting non-invasive cancers early [27]. However, it should be noted that it is very important to detect the non-invasive cancers early before they go on to become invasive. As discussed by Nishikawa in [78], CADe has shown great promise in detecting non-invasive cancers early without significantly decreasing the positive predictive value of a screening program and hence the use of CADe is beneficial to a screening program. Yet, one has to acknowledge that the present day CADe systems are not equally adept at detecting all types of breast lesions. For instance, spiculated lesions (spiculated masses and architectural distortions) are breast lesions that manifest as oriented patterns of converging lines (called spicules) on mammography and carry a very high risk of being invasive [59]. This makes it extremely crucial to detect them early. Unfortunately, clinical studies to evaluate performance of CADe suggest that the rate of detection of invasive cancers such as spiculated lesions is currently not adequate on mammography [94], [4], [11]. The presence of spicules invariably results in a woman being called back for follow-up imaging tests and biopsy. Yet, most CADe systems for spiculated lesions lack tools that attempt to explicitly model and capture the spicules. Most systems also lack visualization aids in the form of spicule annotations that not only would provide strong evidence behind the finding, but would also help the interpreting radiologist determine the exact extent of the disease.

It is important to note that the significance of CADe in the context of present day breast cancer screening cannot be overstated. A major problem facing the radiology community today is that a large number of mammograms are generated due to mammography being performed routinely on many women. The incidence rate of cancer is relatively small: out of every 1000 mammograms that are generated, only to 4-5 mammograms show cancer. The low incidence rate of cancer makes it an extremely challenging task for a radiologist when confronted with a large number of mammograms on a daily basis to correctly detect an existing abnormality [10]. Oversight errors due to the rare occurrence of cancer are often attributed to be one of the main reasons for missed cancer detection [10]. To make matters worse, a study published by Bassett et al. [7] in July 2011 has shown that a number of medical residents have a negative attitude towards interpreting mammograms in future practice. A primary reason for this negative attitude is the stress level due to possible misdiagnosis while interpreting mammograms. This trend is concerning since fewer qualified breast-imaging radiologists are available to read the large number of mammograms that are generated on a daily basis. Another study conducted by Wing and Langlier [105] suggested that the ratio of number of breast imaging radiologists per 100,000 women over the age of 40 years undergoing routine mammographic screening is expected to drop by 17.1% in the period from 2003 to 2025.

Even though CADe systems have been shown to help radiologists detect cancers early, there is still no getting away from the fact that mammography is a 2-D imaging modality. Anatomical noise hinders both human and computer vision in finding cancers. This has lead to recent advances in 3-D breast imaging in the form of breast tomosynthesis [24], stereoscopic (stereo) mammography [31], and breast computed tomography (CT) [12] to augment mammography for early cancer detection. The advent of full field digital mammography, high quality digital displays, and stereoscopic devices have led to rapid progress in these modalities. Breast tomosynthesis has been recently approved by the Food and Drug Administration for clinical use in the U. S., while stereo mammography is currently undergoing clinical trials. Breast CT is still under investigation.

Stereo x-ray imaging and visualization of the breast provides the interpreting radiologist with a 3-D view of the anatomical structures of the breast [31]. This is extremely significant since earlier what was seen on a mammogram was just the superposition of tissue structures from different depth planes, while now with stereo, the tissue structures can be clearly resolved to lie in separate depth planes by a stereo acute observer. Stereo imaging and visualization of the breast has already shown great promise in improving the sensitivity of breast cancer detection and reducing unnecessary patient recalls as demonstrated by the large clinical study conducted by Getty *et al.* [31]. Moreover, stereo visualization has broad applicability than just stereo mammography imaging and has the potential to be used to view projection images acquired during a breast tomosynthesis or breast CT scan.

The advent of stereo mammographic imaging, while still nascent, has opened the door for the development of computational tools for visualizing and interpreting stereo mammograms. These tools will assist the radiologist in interpreting stereo mammograms when the modality is routinely used in the clinical workflow. Example usage of computational tools for stereo mammogram interpretation include measuring lesion properties in 3-D, estimating the depth at which a tumor is located, providing highly localized 3-D information for lesion biopsy, and quantifying parenchymal patterns and breast density in 3-D.

1.2 Dissertation Contributions

This dissertation is divided into two parts. The first part of this dissertation is focused on the problem of spicule annotation on mammography. As described earlier, present-day CADe systems for spiculated masses lack tools that explicitly capture spicules. Our first contribution is an active contour algorithm called *snakules* for the annotation of spicules on mammography [68], [69] and [75]. At each suspect spiculated mass location that has been identified by either a radiologist or a computer-aided detection (CADe) algorithm, we deploy snakules that are converging open-ended active contours also known as snakes. The set of convergent snakules have the ability to deform, grow and adapt to the true spicules in the image via the process of active-contour evolution that optimizes a local matching energy functional. The set of candidate points around the suspect spiculated mass location where snakules are deployed is automatically detected. The algorithm is model-based in that where possible, statistics of prior physical measurements of spiculated masses on mammography are used in the process of detecting the set of candidate points. Observer studies with experienced radiologists to evaluate the performance of snakules demonstrate the potential of the algorithm as an annotation tool that could be used to augment existing spiculated mass CADe systems. We also demonstrate the potential of snakules for extracting spicule features to automatically differentiate between true and false spiculated lesion locations on mammography.

The second part of this dissertation is focused on 3-D breast imaging. Here, we present a novel computational stereo model for estimating a dense disparity map from a pair of stereo mammograms [70]. The problem of reliably estimating a dense disparity map from a pair of stereo mammogram images is very important since this is the first step towards developing computational tools for interpreting stereo mammograms. The proposed stereo model is aimed at preserving isolated impulse or curvilinear masses in the disparity space that arise out of the dense network of fine-scale curvilinear structures often seen on mammograms. Curvilinear structures such as vasculature and spicules are particularly salient structures in the breast, and being able to accurately position them in 3-D is an important goal of this work. Towards solving this difficult aspect of the problem, an important contribution of this work is a novel singularity index that we have developed to reliably detect singular points in images [71], [72], [73]. The singularity index can be configured to detect mass like structures such as impulses in a 1-D signal or curvilinear masses in 2-D images, while rejecting step edges. It can also be configured to do the opposite. The singularity index is inspired by the conditions put forth by Lindeberg [60], [61] and Steger [100] for ridge detection, by Canny's approach to directional edge detection [20], and by an energy operator developed by Teager and studied in detail by Kaiser [44]. The index is a steerable, multi-scale ratio index that employs directional derivatives of gaussians. We present an in-depth theoretical analyses of the index for detecting multi-scale impulses and also study its detection power in the presence of noise. Our analyses and examples on several real world images (mammograms and other naturally occurring images) demonstrate the efficacy of the index for detecting multi-scale curvilinear structures. In order to obtain better estimates of disparity along critical curvilinear structures, the proposed stereo model employs this new singularity index as a constraint in a global energy functional that is minimized to produce a dense disparity map from stereo mammogram images. Experiments on synthetic images with known ground truth and on real stereo mammograms highlight the advantages of the proposed stereo model over the canonical stereo model.

The final contribution of this dissertation is an observer study with experienced breast imaging radiologists, in which we demonstrate the feasibility of viewing breast tomosynthesis projection images stereoscopically [74]. A key differentiating aspect between tomosynthesis projection images and stereo mammogram images is that the x-ray dose under which each tomosynthesis projection image is acquired is much lower than the dose under which a standard mammogram is acquired. Consequently, tomosynthesis projection images have poor contrast due to a low signal to noise ratio when compared to traditional x-ray mammograms. Stereo visualization of tomosynthesis projections has the potential to reveal the 3-D structure of the breast, unlike the current cine or slice-by-slice viewing modes, and could be a viable reading mode in the future. In this study, we evaluated two important questions: 1) Can radiologists perceive 3-D information when viewing the low dose projection images stereoscopically?, and 2) Can mass detection be reliably performed on tomosynthesis projection images using stereoscopic viewing?. The results from our study indicate that reliable stereoscopic 3-D display and interpretation of breast tomosynthesis projection images is possible and there is merit in considering stereoscopic viewing as a viable reading mode for breast tomosynthesis data in the future.

1.3 Dissertation Outline

The remainder of this dissertation is organized as follows. In chapter 2, we review existing work on the computer-aided detection of spiculated lesions on mammography. In chapter 3, we present the snakules algorithm for annotating spicules on mammography. Chapter 4 presents background material on 3-D breast imaging and computational stereo for mammography. In chapter 5, we present our work on the new singularity index for detecting impulse singularities in signals of arbitrary dimensionality. Chapter 6 presents a detailed analysis of the impulse detection power of the index in the presence of noise. We then present our new stereo model that employs the singularity index for better estimating the disparity along curvilinear structures in chapter 7. In chapter 8 we present our observer study in which we demonstrate the feasibility of viewing breast tomosynthesis projection images stereoscopically. In chapter 9, we conclude this dissertation with pointers to interesting areas for future work.

Chapter 2

Computer-aided Detection of Spiculated Lesions on Mammography

2.1 Introduction

Breast lesions can manifest as various findings on mammography – microcalcifications, masses (spiculated and non-spiculated) and architectural distortions. In mammography, a breast mass is defined as a space-occupying lesion that can be seen on at least two projections. Spiculated lesions are characterized by radiating lines or spicules from a central mass region as shown in Fig. 2.1. Spiculated lesions account for about 14% of all biopsied lesions, and about 81% of these are malignant [59]. Hence, spiculated lesions carry a much higher risk of malignancy than calcifications or other types of masses. Spiculated lesions can also form without a central mass region being visible on mammography. Such lesions, referred to as architectural distortions, are also strongly suggestive of malignancy; approximately 48-60% of architectural distortions that are biopsied are found to be cancer [4], [79]. Thus, the early detection of spiculated lesions is critical.

Computer-aided detection (CADe) systems for mammography have been shown to be significantly better at detecting microcalcifications than

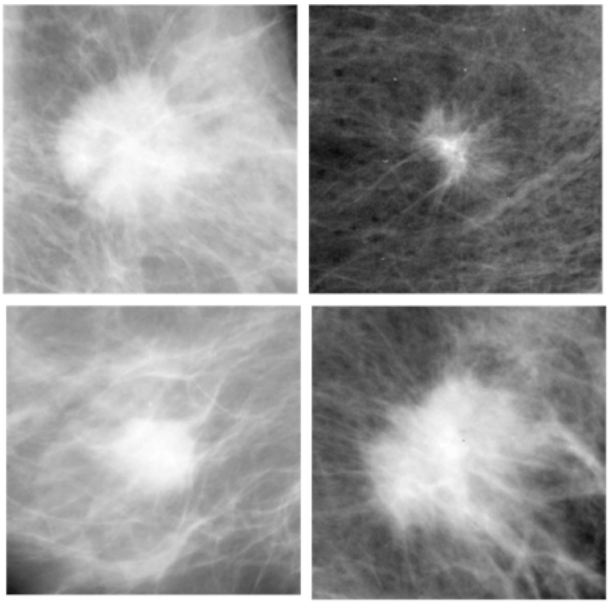

Figure 2.1: Example spiculated masses seen on mammography.

spiculated masses and architectural distortions. Birdwell *et al.* [11] conducted a study in which the characteristics of cancers missed at screening were studied and the ability of CADe to detect these cancers was assessed. Birdwell's study demonstrated that CADe marked most of the cases that were missed by the radiologists, but also showed that CADe systems were better at detecting microcalcifications than masses. Clinical studies to evaluate the performance of commercial CADe systems for microcalcification detection have reported sensitivity close to 100% with false positive per image (FPI) rate being less than 1.0 [51], [92], where as for mass detection the reported sensitivities are only in the range of 67% to 89% with the FPI ranging from 0.40 to 0.74 (e.g., [108], [30], [19], [51], and [92]). Similarly, Baker *et al.* [4] found that for detection of architectural distortion, one commercial system achieved a per-image sensitivity of 38% at 0.70 FPI and another achieved a per-image sensitivity of 21% at 1.27 FPI. Thus, the rate of detection of spiculated lesions is not adequate on standard mammography by either humans or current CADe systems. This has prompted several research groups to develop CADe systems specifically designed to detect spiculated masses [93], [47], [58], [112], [62], [87]. We next review some of the CADe systems for spiculated lesions and discuss their limitations.

2.2 CADe Systems for Spiculated Lesions

Most CADe systems for spiculated lesions are comprised of two stages: a high sensitivity stage to detect suspect lesion locations on the mammogram and a high specificity stage to reduce the number of false positive (FP) candidates that do not correspond to actual lesions. The final outcome of a CADe system is usually a set of marks (also referred to as prompts) on the mammogram identifying suspect lesion locations.

CADe systems designed to detect spiculated lesions usually employ strategies to detect patterns of converging lines. For instance, Karssemeijer and te Brake deployed derivatives of Gaussians at multiple scales to estimate orientations of pixels towards detecting spiculated lesions [47]. The pixel orientation map was used to compute features that were sensitive to patterns of converging lines, and these features were used for classifying suspect locations as possible spiculated lesions. Zwiggelaar *et al.* proposed a statistical approach based on factor analysis to describe oriented patterns of linear structures characteristic of spiculated lesions [112]. The appearance of the central mass was statistically modeled using local-scale orientation signatures extracted from recursive median filtering and approximated using principal component analysis. More recently, Sampat *et al.* [93] proposed a model-based framework for the early detection of spiculated masses on mammography. What makes their approach novel is that they measured physical properties of spiculated lesions on a number of mammograms and developed a statistical model from these measurements. The statistical model served as the basis for determining the parameters of a novel class of pattern matching filters, termed spiculated lesion filters that were deployed to aid in the detection of spiculated masses and architectural distortions. Other methods for detecting spiculated lesions have included the use of directional wavelets and multiresolution analysis of mammograms [58], [62].

2.3 Limitations of Existing CADe Systems for Spiculated Lesions

Most CADe systems for spiculated lesions are developed around the central theme of detecting radial patterns of converging lines. Consequently, most of these methods achieve high sensitivity in detecting suspect spiculated lesion locations. However, a mammogram invariably contains other normal linear structures that are superimposed on one another and resemble a pattern of converging lines. Such locations are routinely marked as suspect locations by the detection algorithms, resulting in a high FP rate. For example, the oriented pattern method proposed by Zwiggelaar *et al.* achieved a sensitivity of 80% at 0.014 FPI for lesions having diameters exceeding 16 mm in a data set containing 29 spiculated masses and 100 normals [112]. However, Zwiggelaar *et al.* reported that the specificity of their orientation pattern method fell below acceptable levels on smaller spiculated lesions [112]. The method proposed by Karssemeijer and te Brake achieved a sensitivity of 80% at 0.4 FPI on a data set containing 19 stellate distortions (spiculated masses and architectural distortions) and 31 normals [47]. However, the attained FPI rate was 1 at a sensitivity of approximately 90% [47]. Likewise, the model-based detection of spiculated lesions proposed by Sampat *et al.* achieved a sensitivity of 80% at approximately 0.75 FPI on a dataset comprising 50 spiculated lesions and 50 normals [93], but the FPI rate increased to 2.7 at a sensitivity of 88%. As pointed out by Zwiggelaar *et al.*, the specificity of CADe systems designed to detect oriented patterns of converging lines could be improved if the radiating spicules could be discriminated from other linear structures in the mammogram [112]. This calls for methods that could accurately capture the profile of spicules and other linear structures. Another problem that exists in the evaluation of these different CADe systems (e.g., [112], [47], and [93]) is that each system has been evaluated on a different, limited dataset of spiculated lesions and it is hard to judge how these systems would perform on a single, large dataset of spiculated lesions and normals. Even clinical studies that have evaluated the performance of commercial CADe systems have found that the detection performance of these systems on stellate distortions is less than acceptable [4]. These numbers highlight the need for methods to improve the overall performance of CADe systems designed to detect spiculated lesions.

Another problem that exists with CADe systems for spiculated lesions is the prompting strategy. As discussed by Astley [3], in the case of spiculated lesions, it is not always clear where to place the symbolic prompt. A symbolic prompt could be placed at the center of the mass, or at the point of convergence (focus) of the spicules [3]. Yet, both these placements of the symbolic prompt might not be optimal in terms of helping a radiologist's efficiency in arriving at a diagnostic decision. CADe systems usually do not explicitly annotate spicules and hence it is left to the radiologist to determine if the symbolic prompt corresponds to a spiculated mass or not. While most radiologists can detect obvious spicules, it would be beneficial if algorithms were developed to detect and explicitly annotate spicules. Such algorithms might help improve clinical productivity.

2.4 The Need for Spicule Annotation Algorithms

Most CADe systems for spiculated lesions lack algorithms that explicitly model and capture spicules. Spicules are salient structures that often influence the interpreting radiologists to call patients back for a diagnostic work-up. Developing algorithms that could explicitly capture and annotate the linear trajectory of spicules would not only provide additional information for classifying suspect spiculated lesion locations on a mammogram, but would also provide rich visual evidence (an example of sophisticated prompting) to the interpreting radiologist and thereby aid in the decision making process. Indeed, other research groups have also tried to develop strategies for automatically detecting and classifying linear structures such as spicules that are seen on a mammogram. For example, Zwiggelaar *et al.* demonstrated the use of cross-sectional intensity profiles as a basis for classifying linear structures seen on a mammogram with particular emphasis on correctly recognizing spicules and ducts [111]. However, Zwiggelaar *et al.* did not explicitly seek to capture the trajectory of linear structures such as spicules; rather, they collected cross-sectional profile information from each linear structure detected on a mammogram using line detection operators and classified them into anatomical types by using a classifier trained on ground truth and cross-sectional information. In another related work, Qian *et al.* have shown that directional wavelets in conjunction with contour tracing can be used to trace spicules explicitly [84]. While the algorithm developed by Qian *et al.* is able to detect spicules of different sizes, it suffers from some drawbacks, most notably - translation/rotational dependence of the decomposition used and ad hoc spicule aggregation rules with lack of model-based design mechanisms.

Chapter 3

Snakules: A Model-Based Active Contour Algorithm for the Annotation of Spicules on Mammography

3.1 Introduction

In this chapter, we present *Snakules* - a model-based active contour algorithm for the annotation of spicules on mammography. As explained in chapter 2, explicitly capturing spicules is important since spicules are salient structures that often suggest the presence of an invasive cancer and hence provide rich visual evidence to the interpreting radiologist to arrive at the appropriate management decision. However, there is a paucity of good annotation algorithms for explicitly capturing spicules and this was the main motivation behind snakules.

At each suspect spiculated mass location that has been identified by either a radiologist or a CADe algorithm, we deploy snakules that consist of converging open-ended active contours also known as snakes [48]. The set of convergent snakules (snakes that seek spicules) have the ability to deform, grow, and adapt to the true spicules in the image, by an attractive process of curve evolution and motion that optimizes the local matching energy. The algorithm is model-based in that statistics of physical measurements of spiculated masses collected from mammograms are used to detect the candidate snakule points from where the snakes originate. We next describe the proposed algorithm in detail.

3.2 Proposed Algorithm

3.2.1 Detecting Candidate Snakule Points

The first step in the snakules algorithm is to detect a set of points on the image from where spicules originate in the region around the suspect spiculated mass location identified by either a CADe algorithm or a radiologist. These candidate points represent locations on the image where snakules would be deployed. The problem of detecting the candidate snakule points can be mathematically stated as the problem of finding the set

$$\begin{aligned} C = \Bigg\{ & (x,y) | (x,y) \in N, |\theta_{x,y} - \psi_{x,y}| < \frac{R}{r_{x,y}}, \psi_{x,y} \in k, \\ & \forall (p,q) \in N, |\theta_{p,q} - \psi_{p,q}| < \frac{R}{r_{p,q}}, \psi_{p,q} \in k,\ r_{x,y} \leq r_{p,q} \Bigg\}, \end{aligned} \tag{3.1}$$

where (x, y) is a candidate snakule point, N represents a neighborhood of pixels under consideration around the suspect spiculated mass location (x_c, y_c), (p, q) represents another point in the neighborhood N, θ is the dominant pixel orientation at a point in N (e.g., $\theta_{x,y}$ is the dominant pixel orientation at (x, y)) , ψ is the direction of a point in N with respect to (x_c, y_c) (e.g., $\psi_{x,y}$ is the direction of the point (x, y) with respect to (x_c, y_c)), R is the radius of a circular disk centered on (x_c, y_c) and towards which the pixel at a given location is directed, r is the Euclidean distance between a point and (x_c, y_c), (e.g., $r_{x,y}$ represents the Euclidean distance between (x, y) and (x_c, y_c)), and k represents the k^{th} orientation bin. The direction ψ at a point, (e.g., (x, y)), is computed as $\psi_{x,y} = tan^{-1}(\frac{y_c - y}{x_c - x})$.

The condition $|\theta_{x,y} - \psi_{x,y}| < \frac{R}{r_{x,y}}$ is the same as defined by Karssemeijer and te Brake [47], in that we consider pixels in a neighborhood around the suspect spiculated mass location that are directed towards a circular disk of radius R centered on the suspect spiculated mass location. However, this condition alone will not suffice to detect the candidate snakule origin points. This condition will yield all the points in the neighborhood N that are directed towards the suspect spiculated mass location. However, we require those pixel points that are not only directed towards the suspect spiculated mass location, but are also closest to the mass. This is captured by the condition on the second line of (3.1), which ensures that of all the pixels that are directed towards the central mass region and whose directions with respect to the suspect spiculated mass location fall in the same orientation bin k, only the point that is closest to the suspect spiculated mass location will be selected. These two conditions yield a set of pixel locations whose dominant orientations are directed towards the central mass region and are closest to the central mass.

The detection of candidate snakule points is carried out on steerable filtered-Radon enhanced regions of interest (ROIs) rather than on the ROIs cropped directly from the mammograms. The primary motivation behind this is to mitigate the effects of noise and clutter caused due to overlapping of out-of-plane tissue structures, which is a common occurrence in mammography due to the projection of 3-D breast structures onto a 2-D image plane. Radon enhancement of spiculated lesions on mammograms is explained in detail in [93]. We briefly review the procedure here.

The Radon transform of a continuous function $f(x, y)$ is defined as

$$g(\rho, \theta) = \int_{y=-\infty}^{y=+\infty} \int_{x=-\infty}^{x=+\infty} f(x, y)\delta(\rho - xcos(\theta) - ysin(\theta)) \, \mathrm{d}x \, \mathrm{d}y, \tag{3.2}$$

where ρ and θ are the parameters of the Radon domain and δ is the Dirac Delta function. Extending the above definition to images, the integral denotes that the value of $g(\rho, \theta)$ for any given (ρ, θ) is the integrated density of the image $f(x, y)$ along the line $\rho - xcos(\theta) - ysin(\theta) = 0$. Thus an n-pixel thick line in the image space $f(x, y)$ maps to n-points located along a column in the Radon domain. It is possible to selectively enhance local peaks comprising a certain number of points in the Radon domain, which corresponds to a linear structure in the image space of the same thickness as the number of points in the Radon domain. The enhancement of local peaks was achieved through a simple linear column filter whose design was based on the average width of the spicules. Once the peaks had been enhanced, we used a Hamming window to attenuate the high frequency noise in the Radon projections (the set of line integrals). Then the inverse Radon transform was computed using the filtered back projection technique to yield a ROI with enhanced linear structures and minimum clutter. Figs. 3.1 shows a ROI extracted from a mammogram depicting a spiculated mass (top), and the corresponding Radon

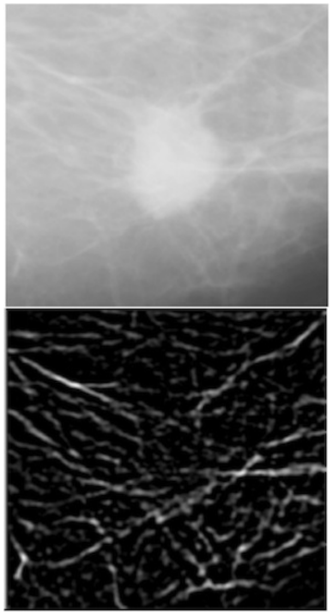

Figure 3.1: Top: a region of interest (ROI) from a mammogram depicting a spiculated mass, Bottom: Radon enhanced ROI depicting enhanced curvilinear structures.

enhanced ROI (bottom). It is evident from Fig. 3.1 (bottom) that the Radon enhanced ROI has curvilinear structures enhanced with suppression of clutter due to overlapping of out-of-plane tissue structures.

The dominant orientation θ at each pixel of the Radon enhanced ROI was computed by filtering the Radon enhanced ROI with a set of steerable quadrature filter pairs comprised of the fourth derivative of a Gaussian and its Hilbert transform [29]. The fourth derivative of a Gaussian and its Hilbert transform provide good angular resolution and can be used for analyzing multiple oriented structures at a single location as described in [29]. The dominant orientation energy response at each pixel of the Radon enhanced ROI was computed as described in [29] and this was used to find the strongest pixel locations in the neighborhood of a suspect spiculated mass location whose dominant orientations are directed towards the central mass region. It is important to note that the steerable filters are applied on the Radon enhanced ROI and not directly in the Radon domain, and the processing is carried out at a single scale.

To test the hypothesis that the estimate of the dominant orientation from the Radon enhanced ROI will be more accurate than from the ROI directly, we simulated linear structures of different widths, with each linear structure oriented at an angle 90 degrees and superimposed on real mammographic backgrounds of different densities. The mammographic backgrounds were ROIs of size 256x256 pixels, and were cropped from a set of 21 digitized mammograms from the Digital Database for Screening Mammography (DDSM) [39], [40]. These mammograms were digitized using the LUMISYS scanner and were used in a prior study by Sampat et al. [95] who collected measurements of physical properties of spiculated masses from experienced radiologists. We ensured that the ROIs were extracted from regions of the mammogram that did not contain the spiculated mass. Each background had one linear structure of length 72 pixels superimposed on it at an angle of 90 degrees. The intensity profile along the length of the structure was set to a constant value equal to the maximum intensity of the mammographic background. This ensured that the linear structure was at least as bright as the background. We considered three different widths (5, 6, and 7 pixels), and each of these linear structures was superimposed on all the 21 backgrounds. The backbone of each linear structure was centered close to the middle column of each mammographic background ROI. Fig. 3.2 illustrates examples of simulated linear structures superimposed on real mammographic backgrounds and the corresponding Radon-enhanced ROIs. The column filter used in the Radon-enhancement was designed to match the width of the linear structure. The dominant orientation was computed from each ROI as well as its Radon enhanced version. The error in the estimation of the dominant orientation was computed at each point along the backbone of the linear structure by comparing the estimated orientation with the ground truth (90 degrees) for both the cases. The error was accumulated across all 21 ROIs for each width of the linear structure. Cumulative distribution curves of the dominant orientation error were generated (Fig. 3.3), and the standard deviation of the error (σ) was computed (Fig. 3.3). The Wilcoxon Sign Rank (WSR) [91] test was also performed to assess if the dominant orientation errors computed from the ROIs and their Radon enhanced versions were significantly different for each of the three different widths of the linear structure. The resulting p-values are also summarized in (Fig. 3.3). These results suggest that the estimate of the dominant orientation is more accurate for different widths of the linear structure when performed on the Radon enhanced ROI rather than on the ROIs cropped directly from the mammograms. We acknowledge that our experimental setup is fairly simple, in that we have considered only a single orientation and a single synthetic linear structure, with no intersections or

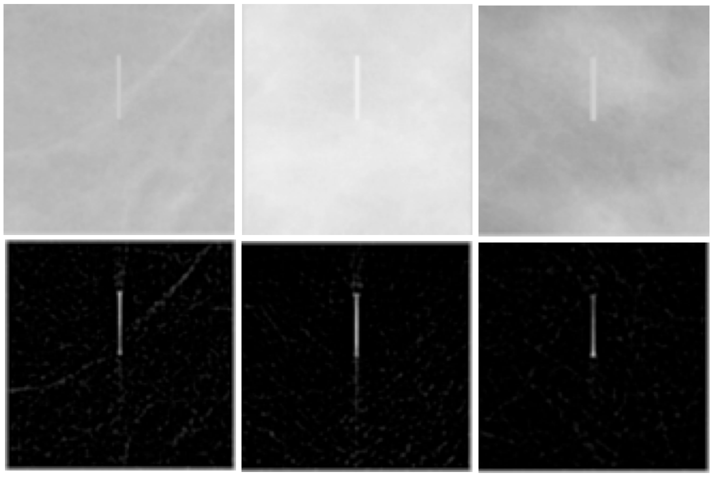

Figure 3.2: Top row: Column 1: 5 pixel wide, Column 2: 6 pixel wide, Column 3: 7 pixel wide simulated linear structures, all oriented at angle 90 degrees and superimposed on real mammographic backgrounds. Bottom row: The corresponding Radon enhanced ROIs.

junctions. However, we sought to evaluate the accuracy in the estimation of dominant orientation on mammographic backgrounds in the simplest possible scenario to help us decide whether to estimate dominant orientations from the ROIs directly or their Radon enhanced versions in the more complex scenario of real data. In Fig. 3.10 (left) and Fig. 3.10 (right), we also show the results of the overall snakules algorithm on an ROI when the dominant orientation was estimated from the Radon enhanced ROI and the ROI cropped directly from the mammogram, respectively. These results also suggest that the quality of the annotations is better when the dominant orientation is estimated from the Radon enhanced ROI rather than from the ROI directly.

The set C of candidate snakule points could contain outliers in the sense that there are points that are farther away from the suspect spiculated mass

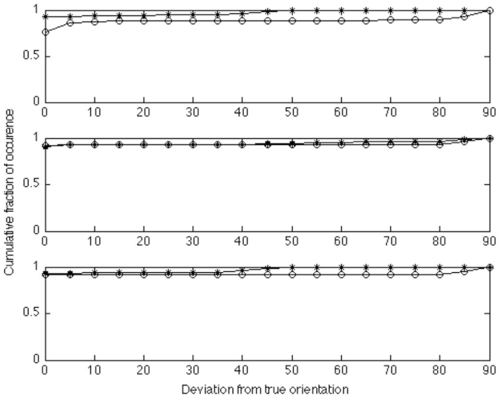

Figure 3.3: Cumulative distributions curves of dominant orientation error (o = ROI, * = Radon enhanced ROI). Top row: 5 pixels wide linear structure (LS), $\sigma = 27.39$ for ROI, $\sigma = 10.60$ for Radon enhanced ROI, Wilcoxon Sign Rank (WSR) p-value < 0.05; Middle row: 6 pixels wide LS, $\sigma = 23.21$ for ROI, $\sigma = 19.17$ for Radon enhanced ROI, WSR p-value < 0.05; and Bottom row: 7 pixels wide LS, $\sigma = 24.18$ for ROI, $\sigma = 11.12$ for Radon enhanced ROI, WSR p-value < 0.05.

location than the other points are. This could arise due to the choice of the heuristic for determining the size of the neighborhood N. To eliminate such outliers, we use a popular distance-based outlier detection technique described in [52]. According to the definition of a distance-based outlier, an object O in a dataset T is an $DB(p, D)$ outlier if at least fraction p of the other objects in the data set lie greater than a distance D from the object O. This is a simple, yet powerful technique [52] to detect outliers in a dataset. In our implementation, D represents the 2-D Euclidean distance of a candidate snakule point from the remaining candidate points and p is set to a value of 50%. In other words, we required at least half the number of remaining candidate snakule points to lie at a distance greater than D for any given point to be considered an outlier.

3.2.2 Snakule Evolution and Growth

Once the candidate snakule points are identified, we deploy open-ended, parametric snakes, originating from these points. A parametric snake or active contour [48] is a parametric curve $v(s) = [x(s)y(s)]^T$ that evolves through the image to minimize the energy functional:

$$E(v(s)) = \int_0^1 \left[\frac{1}{2}(\alpha|v'(s)|^2 + \beta|v''(s)|^2) + E_{ext}(v(s))\right] ds, \tag{3.3}$$

where $v'(s)$, and $v''(s)$ are the first and second derivatives of $v(s)$, representing continuity and curvature of the contour, respectively. The weighting parameters α and β represent the relative importance of the continuity and curvature of the contour. The external energy, $E_{ext}(v(s))$, typically arises from the image and draws the contour towards features of interest such as edges.

A snake that minimizes the net energy has to satisfy the Euler-Lagrange equation $\alpha v'' - \beta v'''' - \nabla E_{ext}(v) = 0$, which can be expressed as the force-balanced equation $F_{int} + F_{ext} = 0$, where $F_{int} = \alpha v'' - \beta v''''$ is the internal force that controls the contours continuity and curvature, and $F_{ext} = -\nabla E_{ext}(v)$ is the external force arising from the image that draws the contours towards edges or lines. Many groups have focused on developing new formulations for the external force (e.g., [107], [57]) and these formulations are primarily geared at improving the capture range of the contour and reducing the impact of image noise on the contour evolution. We use one such formulation for the external force called vector field convolution (VFC) for the evolution of snakules [57]. The VFC force is computed as the convolution of a user-defined vector field kernel with a feature map (e.g., edge map) generated from the image [57]. The vector field kernel is defined such that all vectors in the field point towards the kernel origin [57]. Hence, when the kernel origin is a feature of interest such as an object boundary, then all the vectors in the vector field point towards the object boundary, thereby causing the evolving contour to move towards the object boundary. A significant advantage in using the VFC force as opposed to standard formulations of external forces or more sophisticated formulations such as the gradient vector flow field (GVF) [107] is that the VFC force is robust to spurious edges and noise in the image and provides a large capture range. This is particularly important when evolving snakes on images such as mammograms, due to the large amount of clutter present in these images. Additionally, it is possible to show that the standard edge-based external force and the GVF force can be expressed as special cases of the VFC force [57].

Instead of using a standard edge map as a feature map, we use the Radon enhanced ROI as the feature map. The Radon enhanced ROI has the nice property that the linear structures are enhanced and clutter from spurious edges in the mammogram is reduced, which would otherwise manifest on a traditional edge map. Further, the Radon enhanced ROI has the property that it is non-negative and has a larger value near the enhanced curvilinear structures (Fig. 3.1), and hence these curvilinear structures contribute more to the VFC force than the homogeneous regions of the ROI. This property will cause the VFC external force to attract the evolving contour close to the curvilinear structures of interest. The VFC force replaces the standard external force $F_{ext} = -\nabla E_{ext}(v)$ in the force balance equation $\alpha v'' - \beta v'''' + F_{vfc}(v) = 0$. The snake update equations remain the same and are as given in [57].

We adopted an approach that lets the snakules grow and subsequently evolve to trace the spicules. This was motivated by the fact that the true length of a spicule is not known beforehand. Dropping snakes much longer than the actual spicule brings about instability as the snakes evolve towards structures that are not spicules. The idea of growing snakes was first described by Berger [9] as a robust alternative to dropping long snakes to trace open-ended curvilinear structures in images and we adopted a similar strategy. The initial snakule segment is deployed as a short straight line of length 10 pixels at the detected candidate snakule point. This choice of snakule segment length was arbitrary, and the only consideration was to prevent instability in the snake evolution process by keeping the segment relatively short. The initial orientation of this straight line is estimated from the dominant pixel orientation θ computed from steerable filter analysis. This short snakule segment then deforms towards the enhanced spicule under the influence of the VFC force. Once the deformation is complete, the deformed curve is extended in the tangent direction by introducing another short snakule segment, which is a straight line of length 10 pixels. The new segment then deforms and the iterative process of snakule growth and deformation continues until a stopping criterion is met. We used a curvature-based stopping criterion, in which the growth of a snakule was stopped at a point where the curvature of the snake exceeded a 30 degrees limit. We used a measure of curvature described in [104] that is bounded, with values in the interval $[0\ \pi]$. The curvature limit was based on excitatory-inhibitory models of human perception of contour grouping [109]. Fig. 3.4 illustrates three iterations of a growing snakule. It is important to note that our strategy of growing snakules is similar to how a radiologist would annotate a spicule on a mammogram using an electronic interface such as a stylus or a mouse. Finally, once the evolution and the growth of a snakule has stopped, we perform a simple post-processing operation to check if the current snakule has annotated a linear structure that has already been annotated by any of the other snakules. This could happen if two snakules originate from relatively close points, and end up being attracted towards the same linear structure. To prevent this redundancy, at every point

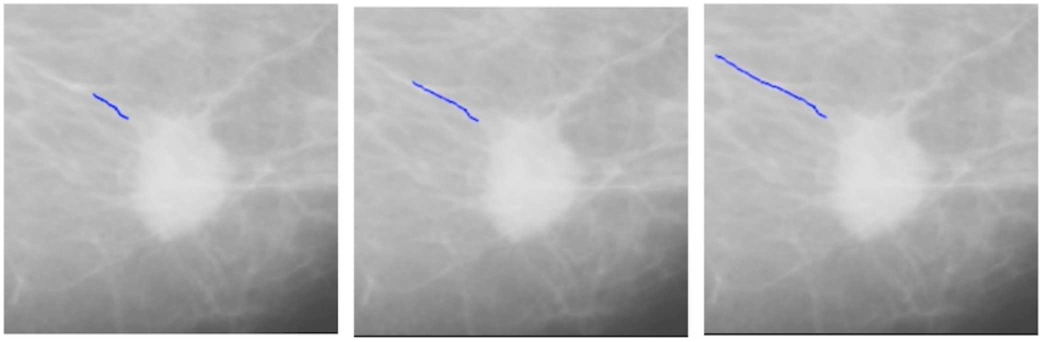

Figure 3.4: Three iterations of a growing snakule illustrating the growth of a snakule until the curvature-based stopping criterion is met.

on the current snakule, we check if there is at least one point in any of the other snakules that lies within a square window of size 3 × 3 pixels centered at the point on the current snakule. If there exists one such point in the 3 × 3 window, then the current snakule is discarded. The size of the window is kept deliberately small and we end up only discarding snakules that are very close to any of the other snakules.

3.3 Experimental Methodology

3.3.1 Data set

The data set for experimental evaluation was comprised of 31 spiculated masses identified on mediolateral oblique (MLO) views that were randomly selected from The Digital Database for Screening Mammography (DDSM) [39], [40]. Cases from two different scanners, LUMISYS (scanner resolution 50 micrometers per pixel) and HOWTEK (scanner resolution 43.5 micrometers per pixel), and a range of density and subtlety ratings were represented in this

No. of cases per scanner		No. of cases per density rating				No. of cases per subtlety rating				
LUMISYS	HOWTEK	1	2	3	4	1	2	3	4	5
16	15	8	8	7	8	1	3	7	7	13

Table 3.1: Statistics of the data set from the digital database for screening mammography (DDSM) used for evaluating snakule annotations. Subtlety of lesion appearance ratings were obtained from DDSM, where subtle = 1, and obvious = 5. The density ratings, which indicate the breast fibroglandular tissue density were collected from an experienced radiologist. Density = 1 indicates primarily fatty tissue, density = 2 indicates scattered fibroglandular tissue, density = 3 indicates heterogenously dense tissue, and density = 4 indicates extremely dense tissue.

cohort. Table 3.1 summarizes the statistics of the dataset. Regions of interest (ROIs) were extracted from each case and the experiments were conducted on the ROIs. Each ROI was defined such that the central mass and all the spicules were clearly included. The 31 ROIs were broken down into 2 sets - a set of 11 ROIs (denoted as development set in Table 3.1) was used for the development and validation of the snakules algorithm and a set of 20 ROIs was reserved exclusively for the evaluation of the algorithm (denoted evaluation set in Table 3.1). The ROIs on which snakules were deployed were sub-sampled by a factor of four (along both the row and the column dimension). This was done to reduce the computation time.

3.3.2 Parameter Design

The choice of the parameter values used in the detection of the candidate snakule points, where possible, was based on prior measurements of physical properties of spiculated masses. These measurements were obtained as a part of a study conducted by Sampat et al. on a set 21 digitized mammograms (LUMISYS scanner) from DDSM [95]. Note that this was a separate set of 21 cases and there was no overlap of cases between this set and the set used for the evaluation of snakules. ROIs containing spiculated masses were extracted from these mammograms and measurements of their physical properties were collected from two experienced radiologists. The physical properties measured included the major axis of the central mass region, the width of a spicule, length of a spicule, and the number of spicules around the central mass. Sampat et al. showed that experienced radiologists can reliably measure these physical properties of spiculated masses with a high degree of inter observer agreement [95].

The design of the linear column filter used to create the Radon enhanced ROI was based on the average spicule width. We conducted an analysis of the distribution of the spicule width (in pixels) measured by the radiologists [95]. We found that the spicule width measurement followed a log-normal distribution as illustrated in Fig. 3.5. We used the measurements made by the more senior radiologist for choosing the parameter values. The average spicule width was found to be 968.42 ± 615.28 micrometers from the spicule width distribution of the more senior radiologist (Fig. 3.5, left). This value of average spicule width translated to 5 pixels for sub-sampled ROIs extracted from mammograms scanned using LUMISYS and 6 pixels for sub-sampled ROIs extracted from mammograms scanned using HOWTEK. We chose to set the parameter value for the average spicule width to be equal to 6 pixels since the snakules algorithm was to be deployed on cases from both the scanners.

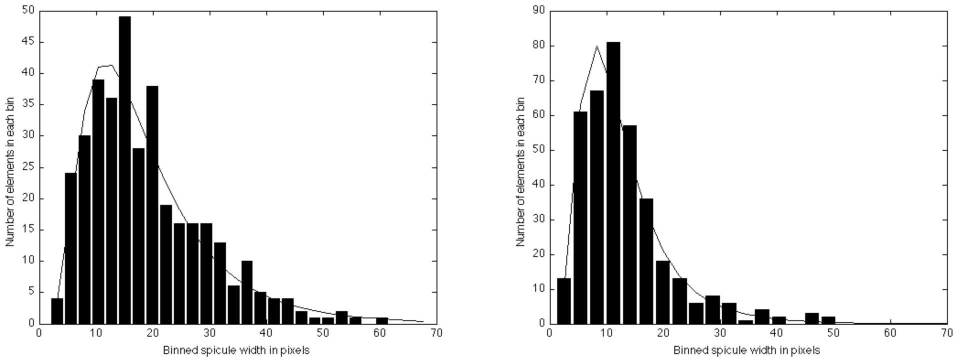

Figure 3.5: Distribution of the spicule width measurement. Left: average spicule width = 968.42 ± 615.28 micrometers, mean squared error (MSE) of log-normal fit = 0.65, Right: average spicule width = 641.25 ± 392.69 micrometers, MSE of log-normal fit = 0.29.

Consequently, we used a column filter $[-3\ 1\ 1\ 1\ 1\ 1\ 1\ \ -1.5\ \ -1.5]^T$ to enhance local peaks that are comprised of 6 points in the Radon domain. It is important to note that since the filter coefficients sum to 0, the response of the filter to slowly varying values in the Radon domain will be close to 0 or very small.

The choice of the parameter R, the radius of the circular disk as defined in (3.1), was based on physical measurements of the major axis of 21 spiculated masses. The median radius of a spiculated mass was found to be 6.9 ± 3.98 mm from the measurements of the major axis performed by the more senior radiologist. This value of median radius translated to 35 pixels for sub-sampled ROIs extracted from mammograms scanned using LUMISYS and 40 pixels for sub-sampled ROIs extracted from mammograms scanned using HOWTEK. We chose to set the parameter R to be equal to 40 pixels since the snakules algorithm was to be deployed on cases from both the scanners.

The space around the suspect spiculated mass location (x_c, y_c) as defined in (3.1) was divided into 24 orientation bins with a bin width of 15 degrees. From the measurements of the physical properties of spiculated masses, we found the average number of spicules to be equal to 18 ± 6. Our choice for the number of orientation bins was based on this finding from the measurements of physical properties of spiculated masses. Finally, the neighborhood N in (3.1) was defined as a circular torus whose inner radius was set to a value of 31 pixels and the outer radius was set to a value of 99 pixels. These numbers, though somewhat arbitrary, were set to account for the variation in the mass sizes of different spiculated masses and the varying lengths of their spicules. Fig. 3.6 illustrates the detection of candidate snakule points with the parameters set as described in this section.

The values the parameters α and β in (3.3) have been set to 0 and 1, respectively. A zero value of α at a point along the snake could result in a discontinuity at that point. However, since each snakule segment that deforms is a relatively short segment (10 pixels), discontinuities are rare. A value of 1 for β places an emphasis on the segment being free of corners, which is a desirable property for tracking curvilinear segments.

3.3.3 Evaluation of Snakule Annotation

To evaluate the performance of snakules for annotating spicules, we conducted an observer study with three experienced radiologists who judged

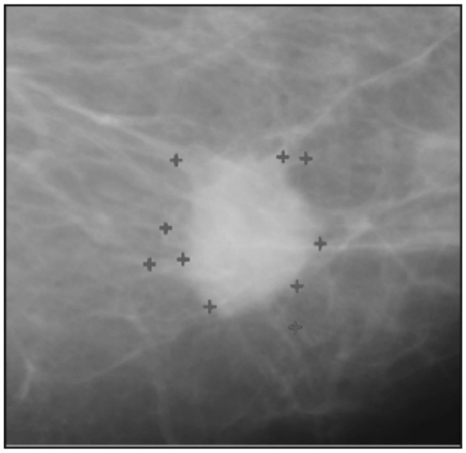

Figure 3.6: Detection of candidate snakule points (each candidate point identified on the ROI is denoted by a +).

the quality of the annotations. Two randomly chosen sets of 10 distinct spiculated masses were used for this study. Each set consisted of ROIs manually centered on a spiculated mass that was cropped from the mammogram, and neither of these sets had been previously used during the development and the initial qualitative validation of the algorithm. The evaluation of the snakules algorithm was performed in two phases. In the first phase, the snakules algorithm was used to annotate one set of ROIs, while the other set was annotated by an experienced, non-clinical observer (self) using the NIH ImageJ freehand annotation tool. The images were not edited prior to annotating using the free hand annotation tool and were annotated sequentially. No magnification was used on any of the images and the annotations were performed in one sitting. The two sets of annotations were then presented to the three radiologists in a randomized order. The three radiologists had a collective experience of more than 40 years in interpreting mammograms. In the second phase, which was conducted six months later, the set of 10 ROIs that were previously annotated using snakules, were manually annotated by the non-clinical observer (self), and the set of ROIs that were previously manually annotated were annotated using snakules. The two sets of ROIs were again presented to the three radiologists in a randomized order. Having the two phases six months apart ensured that the radiologists did not remember the cases, and the annotations. Thus, at the end of the two-phase study, there were 20 ROIs, each of which was annotated using both the methods (snakules and manual). The two phases of the study were blinded in that the radiologists did not know anything about the techniques used to annotate the ROIs, and the radiologists were not allowed to discuss with one another. The main goal of the study was to judge if the quality of the annotations performed by the snakules method on a set of spiculated masses was equivalent to the quality of annotations performed by an experienced observer.

In each phase, the radiologists were presented with a questionnaire in which they were asked to report the following quantities for each of the 20 annotations that were presented in a randomized order: 1) The number of linear structures that were annotated but did not correspond to actual spicules (i.e., the number of false positive (FP) annotations) 2) The number of obvious spicules that were missed (i.e., the number of false negatives (FN)) 3) Of the obvious spicules correctly annotated (in other words, the true positives (TP)), the number of spicules that had their lengths correctly identified. The radiologists were specifically instructed to focus on obvious spicules, since it is nearly an impossible task for either a human or a computer vision algorithm to identify all the spicules around a spiculated mass.

For each radiologist, we computed the following measures from the three quantities reported by them:

$$
\begin{aligned}
P &= \frac{\#TP}{(\#TP + \#FP)} \\
R &= \frac{\#TP}{(\#TP + \#FN)} \\
L &= \frac{(\#TP\ whose\ length\ was\ correctly\ identified)}{\#TP},
\end{aligned}
\tag{3.4}
$$

where P is the precision of the annotation method defined as the fraction of all the annotations accumulated from the 20 cases that correspond to a true spicule, R is the recall of the annotation method defined as the fraction of obvious spicules correctly identified and annotated across all 20 cases, and L is the fraction of correctly annotated spicules whose length was correctly identified across all 20 cases by the annotation method (in other words, the accuracy of correctly identifying the true spicule length). For a good annotation method, all the three measures described above in (3.4) should be close to 1.

An equivalence test for binomial random variables [5] was performed on each of the three measures described above in (3.4) to statistically assess the quality of annotations performed by the snakules method against the quality of annotations performed by the manual method. An equivalence test was the appropriate choice since we were primarily interested in establishing equivalence of the two methods rather than the difference between the two methods.

We made a few assumptions regarding our experimental design in order to perform the equivalence test for binomial random variables. Firstly, for assessing the equivalence of precision score of the two methods, we made an assumption that each annotation in the set of all annotations accumulated from the 20 ROIs was an independent and identically distributed (i.i.d) Bernoulli trial. The outcome of each trial was a true spicule (success) or a false spicule (failure) as judged by the radiologist. Secondly, for assessing the equivalence of recall score of the two methods, we made an assumption that each spicule in the set of all obvious spicules (spicules that were evident to a radiologist) that were accumulated from the 20 ROIs was an i.i.d Bernoulli trial. The outcome of each trial was a successful annotation (success) or a missed annotation (failure). Finally, a similar assumption of a set of i.i.d Bernoulli trials was made to statistically assess the equivalence of the accuracy of length of true spicules identified by the two methods. We acknowledge that these were simplifying assumptions as the detection of a spicule from a dense mammographic background might not be an i.i.d Bernoulli trial and might depend on factors such as presence of other spicules, and the mammographic background in itself. Given these assumptions, it was possible to express the three performance measures of the two methods as probabilities of success of two binomial random variables and the effectiveness of the snakules method was judged by an equivalence test for binomial random variables.

For testing equivalence of the quality of annotations performed by snakules against the quality of annotations performed by an experienced observer, the null hypothesis is that the quality of annotations performed by snakules is not equivalent to the quality of annotations performed by an experienced observer. The alternate hypothesis is that the quality of annotations performed by the two methods are equivalent to one another. The null and alternate hypothesis can be mathematically defined as [5]:

$$\begin{aligned} H_0 &= |p_s - p_m| > \Delta \\ H_1 &= |p_s - p_m| < \Delta, \end{aligned} \tag{3.5}$$

where H_0 and H_1 denote the null and alternate hypotheses, respectively, p_s, and p_m represent the population binomial probabilities of success (one of the three measures described in (3.4)) of the snakules and manual methods respectively, and Δ is a positive quantity specified by the experimenter. A confidence interval approach was used to decide if the null hypothesis could be rejected [5]. There are many variants of the confidence interval approach and these are based on considerations of prior knowledge of the population binomial probabilities, sample size and statistical power (the probability of finding a statistically significant result when there exists one) [5]. We used the continuity corrected simple asymptotic interval (SAIC) to decide if the null hypothesis of no equivalence has to be rejected or not. According to SAIC [5], the null hypothesis H_0 is rejected if

$$\begin{aligned} \hat{p_s} - \hat{p_m} \pm Z_\alpha \Bigg[\Bigg\{ \frac{\hat{p_s}\left(1-\hat{p_s}\right)}{n_s} \\ + \frac{\hat{p_m}\left(1-\hat{p_m}\right)}{n_m} + \frac{\left(\frac{1}{n_s}+\frac{1}{n_m}\right)}{2} \Bigg\} \Bigg]^{0.5} \in [-\Delta \ \Delta]. \end{aligned} \tag{3.6}$$

In (3.6), p_s, p_m, and n_s, n_m are the sample binomial probabilities of success and the sample sizes of the snakules and the manual annotation methods respectively, Z_α is $(1-\alpha)100^{th}$ percentile of the standard normal distribution, and α is the nominal significance level. We used the SAIC confidence interval since when there is no prior knowledge of the population probabilities and when the sample sizes are moderate (50 - a few hundred), the SAIC test runs a lesser risk of being anti-conservative, and offers reasonable statistical power. The decision to use SAIC was based on the recommendations made in [5]. In our analysis, the value of α was set to 0.05 and we considered three values for Δ: 0.25, 0.35, and 0.45. The choice of the parameter Δ affects the outcome of the test. For example, if Δ were set to 0.25, then the snakule and manual annotation methods are considered equivalent if their probabilities of success are within 0.25 of each other, and not equivalent for a larger difference in the probabilities. Consequently, a smaller value for Δ makes the interval tighter and represents a strict criterion for rejecting the null hypothesis of no equivalence, while a larger value for Δ represents a more lenient criterion for rejecting the null hypothesis. The three values we used for Δ, i.e., 0.25, 0.35 and 0.45 represented a strict, moderately lenient, and lenient criterion, respectively. For assessing inter- and intra-observer variability in clinical tasks such as radiographic measurements, a Δ value of 0.25 is considered to be a strict criterion, while a value of 0.30 is considered to be a more lenient criterion, and if the variability is within these values, it is considered acceptable [95]. However, to the best of our knowledge, no prior literature exists on comparing effectiveness of spicule annotation algorithms, and hence we used three values for Δ to represent a strict, moderately lenient, and lenient criterion. Further, our choice of Δ might seem liberal, for instance if $\Delta = 0.45$ and the two methods turned out to be equivalent, then the methods are equivalent even if their probabilities of success are within 0.45 of each other. However, it is important to note that we are comparing the annotations performed by an automated algorithm against those performed by an experienced observer. The manual annotations made by an experienced observer and evaluated by three radiologists can be treated as the gold standard and they provide a high bar for assessing the effectiveness of an automated algorithm.

3.3.4 Snakules for Classification

We also conducted a preliminary experiment in which we explored the use of snakules for extracting spicule features and their subsequent use for automatic classification of candidate spiculated mass locations on mammography. The dataset for the classification experiment consisted of 36 MLO view spiculated mass mammograms (LUMISYS) retrieved from DDSM [39], [40]. A CADe algorithm developed by Sampat et al. [93] was used to screen these mammograms and a set of true lesion and false positive locations for each mammogram was output by the algorithm. Out of the true lesion locations, we only considered those locations deemed as most probable lesion locations by the CADe algorithm. An ROI was centered on each of these locations and was cropped from the mammogram. The final dataset consisted of a total of

312 ROIs with 36 positive instances and 276 negative instances.

Snakules were deployed on each ROI and the following features were extracted- 1) Average Snakule Contrast (f_1): We define snakule contrast C_S as $C_S = \sqrt{\frac{1}{N_S-1}\sum_{\forall(x,y)\in S}(I_{(x,y)} - \mu)^2}$, where S denotes a snakule, $(x, y) \in S$ denotes the coordinates of a point through which the snakule passes, $I_{(x,y)}$ denotes the interpolated image intensity at the point (x, y), N_S is the number of points through which the snakule passes, and μ is the average interpolated image intensity of the snakule trajectory (I_S) and two background trajectories (I_{B_1} and I_{B_2}). The two background trajectories are computed at a distance marginally greater than (by approximately 4 pixels) half the width of the structure annotated by the snakule in the directions of the inward and the outward normal at every point on the snakule. Fig. 3.7 illustrates the snakule and background trajectories (left). The average snakule contrast is computed as the average contrast over all snakules identified on the ROI. The intuition behind using average snakule contrast as a feature is that visually perceivable spicules are normally perceived as bright structures with good contrast relative to the background. 2) Median Distance from Point of Convergence (f_2): Convergence of spicules is a characteristic feature of spiculated masses. We estimated a point of convergence (P_c) of spicules in the ROI space and subsequently computed the median distance of all linear structures annotated by snakules from P_c. Our hypothesis is that when an ROI actually contains a spiculated mass, then the median distance of linear structures from the point of convergence will be lower than the median distance from the point of convergence of linear structures annotated on an ROI not containing a spiculated mass. The problem of finding P_c can be posed as a search problem in which a point that is at a minimum median distance from all linear structures annotated in the ROI space is sought. We employ a greedy search strategy over the entire ROI space to find the point of convergence. Essentially, for every integer pixel location of the ROI, we compute the distance between the pixel location and a linear fit of each snakule. The point whose median distance from all lines is minimum is designated as the point of convergence and the minimum median distance is used as a feature for classification. It is important to note that the point of convergence can be computed when there are at least two linear structures annotated by snakules. If this is not the case, then the median distance is assigned a bookkeeping value of -1 for that particular ROI. Fig. 3.7 illustrates the point of convergence (denoted by a red +) on a spiculated mass ROI annotated using snakules (right). 3) Histogram of Normalized Squared-Intensity Deviation (f_3): As our third feature, we constructed a histogram of the normalized squared-deviation of the interpolated image intensity from the average intensity (defined earlier) at every point on the trajectory of the snakule. For every snakule, the squared-intensity deviation is normalized with respect to the maximum squared-intensity deviation for that snakule. We built the histogram using the k-means vector quantization technique (VQ) in which the centers of the VQ codebook were computed using k-means clustering of squared-intensity deviation values accumulated from the training set of ROIs containing both positive (ROIs with spiculated mass) and negative (ROIs with

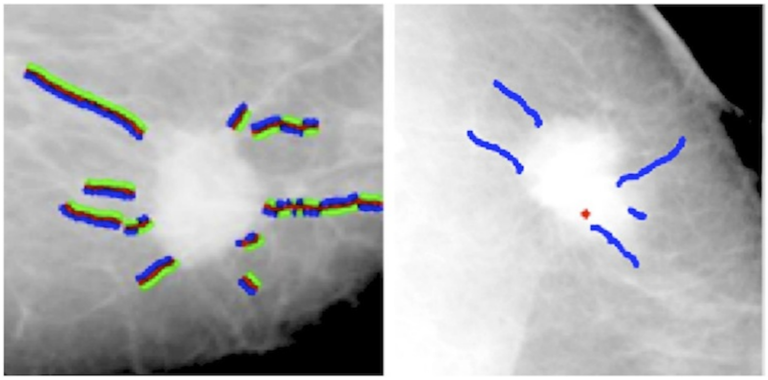

Figure 3.7: Spicule feature extraction using snakules: Left: snakule trajectory (red) and background trajectories (blue and green) on a spiculated mass ROI, Right: Point of convergence (red +) illustrated on a spiculated mass ROI.

no spiculated mass) instances. The number of k-means clusters to build the histogram was set to 7.

A two-fold cross-validation experiment was performed in which 18 randomly chosen mammograms were used for training the classifier and the remaining 18 mammograms for testing the classifier. This resulted in two sets of ROIs: one comprising of 149 instances (18 positive instances, 131 negative instances) and the other comprising of 163 instances (18 positive instances, 145 negative instances). The classifier used was logistic regression. Since the ratio of the number of positive instances to the number of negative instances contained in the training set in each fold was lower than 0.5, we randomly sampled the positive instances with replacement to increase the ratio to 0.5 prior to training the classifier. The classifier was then tested using the features extracted from each ROI in the testing set and the performance of the classifier was evaluated using the area under curve (AUC) metric of the Receiver

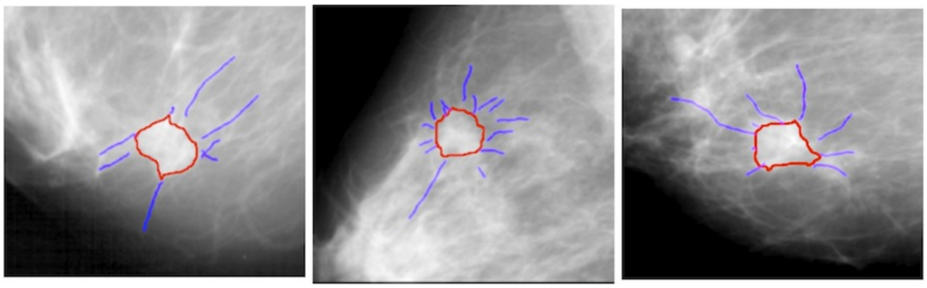

Figure 3.8: Examples of ROIs annotated using snakules (central mass annotated manually by the author).

Operating Characteristic (ROC) curve.

3.4 Results

3.4.1 Annotation Experiment

Fig. 3.8 and Fig. 3.10 (left) show examples of ROIs annotated using snakules, while Fig. 3.9 shows examples of ROIs annotated manually. Qualitatively speaking, the snakule annotations look encouraging on the ROIs illustrated in Fig. 3.8 and Fig. 3.10 (left) . The snakule annotations also suggest that the method is fairly robust to variation in the size of the central mass. Fig. 3.11 illustrates the variation in annotations with the initial location (x_c, y_c) as defined in (3.1). The choice of initial location does affect the quality of annotations by generating some false positives, but as evident in Fig. 3.11, the prominent spicules are usually annotated.

Table 3.2 summarizes the three performance measures described in 3.4 and the results of the statistical tests for equivalence. The following conclusions can be made from Table 3.2: 1) Firstly, when an obvious spicule is present

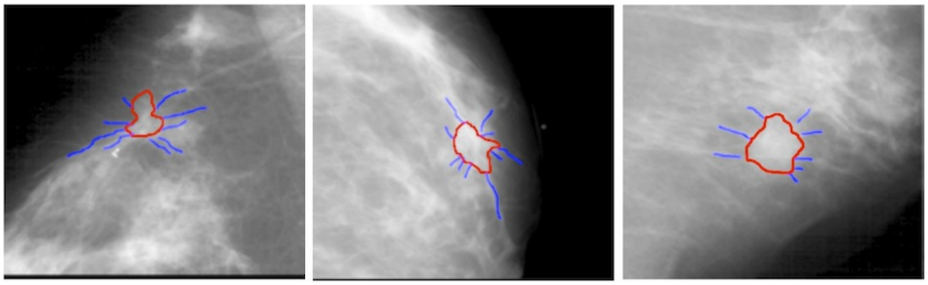

Figure 3.9: Examples of ROIs annotated manually by the author.

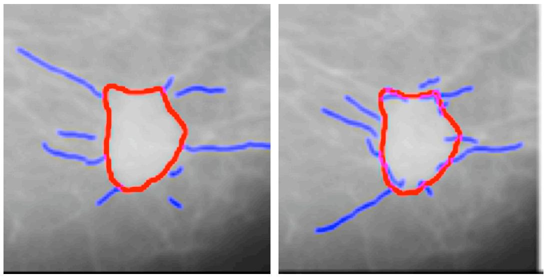

Figure 3.10: Left: snakule annotations when dominant orientation is computed from the Radon enhanced ROI, Right: Snakule annotations when dominant orientation is computed from the ROI cropped directly from the mammogram. Central mass was annotated manually by the author.

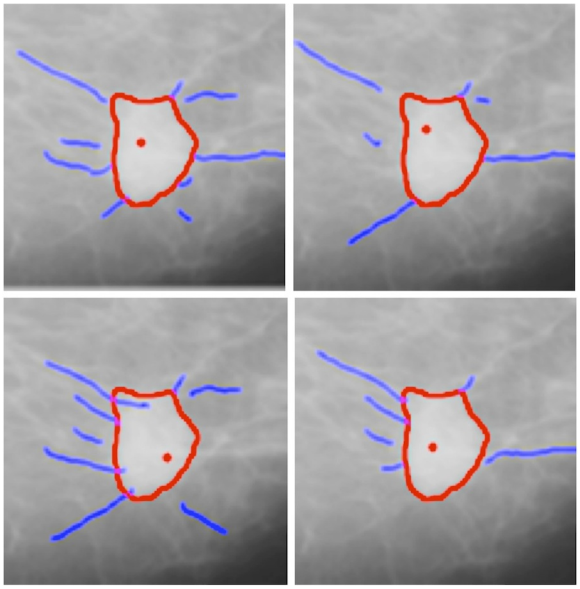

Figure 3.11: Variation in snakule annotations with the initial location indicated by a + (central mass annotated manually by the author).

on the mammogram, the snakules method annotates the spicule with a high probability of success as evident from the recall scores (maximum recall = 0.84). This was consistent across the evaluations by all the three radiologists. Further, the recall scores of the snakule and manual annotations were found to be statistically equivalent at a strict criterion ($\Delta = 0.25$) for all the three radiologists, which was encouraging. 2) Secondly, as evident from Table 3.2, the snakules method annotates spicules with a reasonably good precision (maximum precision = 0.76) and this was again consistent across the evaluations by all the three radiologists. When compared to the recall, the precision of the snakules method is lower, suggesting that the method annotates a few curvilinear structures such as blood vessels that are directed towards the central mass region. Fig. 3.12 illustrates an example of one such ROI in which the radiologists thought that blood vessels had been annotated. The precision scores of the snakules and manual methods were found to be statistically equivalent at a lenient criterion (Δ= 0.38, Δ= 0.45 is reported in Table II) for all the three radiologists. 3) Finally, the maximum accuracy of correctly identifying the true spicule length was 0.95 for the snakules method. However, the length accuracy value was most variable across the three radiologists (0.50, 0.79, and 0.95). Further, for one radiologist, the length accuracy scores of the snakules and manual methods were found to be statistically equivalent at a strict criterion ($\Delta = 0.25$), while for another radiologist the length accuracy scores were found to be statistically equivalent at Δ= 0.26 (Δ= 0.35 is reported in Table 3.2). However, for the third radiologist, the length accuracy scores of the two methods were not found to be statistically equivalent at any of the three chosen Δ values. This suggests that while the radiologists are fairly consistent in identifying the obvious spicules and the false positive annotations, they are not very consistent in identifying the true length of the spicule. However, for two out of the three radiologists, the length accuracy scores of the snakules and manual methods were found to be statistically equivalent at Δ=0.26, which was encouraging from the point of view of extracting meaningful physical properties of spicules such as length that can subsequently be used as a feature in the classification of linear structures evident on a mammogram.

Perf. Measure	Radiologist	Probabilities		Sample size		Null hypothesis rejected?		
		$\hat{p_s}$	$\hat{p_m}$	n_s	n_m	$\Delta = 0.25$	$\Delta = 0.35$	$\Delta = 0.45$
Precision	1	0.69	0.90	187	163	No	No	Yes
	2	0.76	0.97	187	163	No	No	Yes
	3	0.71	0.90	187	163	No	No	Yes
Recall	1	0.81	0.87	160	169	Yes	Yes	Yes
	2	0.84	0.88	169	180	Yes	Yes	Yes
	3	0.80	0.83	166	176	Yes	Yes	Yes
Length acc.	1	0.50	0.84	129	147	No	No	No
	2	0.95	0.98	142	158	Yes	Yes	Yes
	3	0.79	0.86	133	146	No	Yes	Yes

Table 3.2: Precision, recall, and spicule length accuracy scores computed from radiologists evaluations of snakules and manual annotations. Also shown are the conclusions of the statistical test for equivalence for all three measures.

Figure 3.12: Example of an ROI in which the snakules method has annotated quite a few curvilinear structures that are non-spicules but directed towards the central mass.

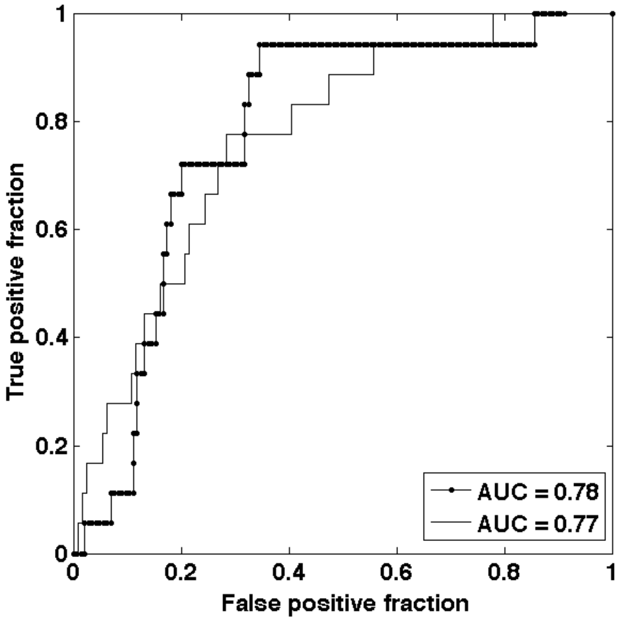

Figure 3.13: ROC curves of the two-fold cross validation experiment.

3.4.2 Classification Experiment

Fig. 3.13 illustrates the corresponding ROC curves obtained for the classification experiment. The AUC values of the two folds were 0.78 and 0.77. The results of the classification experiment are promising as they suggest that just by using only spicule features extracted from snakules, there is enough discriminatory power to differentiate between true lesion and false lesion locations on mammography.

3.5 Summary

In this chapter, we have presented a novel algorithm termed snakules for the annotation of spicules on mammography. The algorithm is model-based in that its design is guided by statistics of physical measurements of real spiculated masses on mammography. Certain generic aspects of the design such as the use of Radon transform to enhance linear structures and compute dominant orientation, and the VFC force field computed from the Radon enhanced ROI for active contour deformation could be useful for other medical image segmentation tasks. We foresee snakules being useful for a variety of tasks. The results of our observer study demonstrate the reliability of snakules in annotating obvious spicules. This suggests that snakules could be used as an additional prompting tool by the radiologists to visualize the spicules and determine the extent of the disease. Snakules could also be used in conjunction with online annotation tools (e.g., [110]) to annotate spiculated masses on mammograms that can be used for training radiology personnel (such as residents). Finally, snakules could be useful for extracting spicule features, which can be subsequently used to distinguish true spiculated lesion locations from non-lesion locations on a mammogram and improve the specificity of CADe systems. The results from our classification experiment are promising in this regard.

Even though radiologists and CADe systems for spiculated lesions could benefit from tools such as snakules, their performance is constrained by the fact that mammography is a 2-D imaging modality. As explained before, anatomical noise is a problem that hinders both humans and computer-based systems in finding early cancers. This has resulted in the radiology community exploring alternate and adjuvant 3-D breast imaging modalities for early breast cancer detection and diagnosis. In the following chapter, we will review some of the recent developments in 3-D breast imaging.

Chapter 4

Stereo Imaging and Visualization of the Breast

4.1 Introduction

One of the major developments in radiology over the last 15 years has been the advent of new 3-D x-ray based breast imaging modalities such as digital breast tomosynthesis [76], digital breast CT [12], [13], and stereo mammography [31]. These developments are significant as they promise to alleviate the effects of a key limitation of mammography - anatomical noise caused by overlapping of out of plane tissue structures on mammographic projections of the breast. As described earlier in Chapter 1, anatomical noise often hinders accurate detection and diagnosis of breast lesions in mammography. The problem posed by anatomical noise is exacerbated in mammograms of women with dense breasts. Even though computer-aided detection systems have helped radiologists in interpreting mammograms, anatomical noise affects both human and computer vision. This major limitation of mammography has spurred the radiology community to develop alternate and adjuvant 3-D x-ray based breast imaging modalities for early detection of breast cancer. In this chapter, we provide a brief review of stereo mammography. We also discuss computational stereo for mammography and review relevant work on disparity estimation.

4.2 Stereo Mammography

Stereo mammography has been made possible by the advent of full field digital mammography, high quality digital displays, and stereoscopic devices. In a stereoscopic setting, a radiologist is provided with a 3-D view of the breast that enables the radiologist to see the lesion in 3-D separated from the overlying and underlying tissues. The stereo mammogram is acquired as two x-ray projection images of the breast taken at two slightly different angles. The angle of separation between the two x-ray images is typically between 4-10 degrees. The breast and the detector remain fixed in position while the x-ray source is rotating. The net radiation dose required for the stereo acquisition is kept equivalent to the dose required for digital mammography by distributing the dose between the two projections.

The resulting stereo mammogram is then viewed using a stereoscopic display and cross-polarized lenses. Stereo acute observers can fuse the stereo mammogram pair and perceive the structures throughout the breast in 3-D. Stereo imaging contrasts with breast tomosynthesis and breast CT in which multiple projection images are acquired over a much wider angular range (15-50 degrees for tomosynthesis and 360 degrees for breast CT) and 3-D volumetric data is reconstructed. However, as demonstrated in [103] and in our own study [74], which we describe in detail in Chapter 8, it is possible to view tomosynthesis projection images using a stereoscopic display. Stereo imaging and visualization of the breast has already shown great promise in improving the sensitivity of breast cancer detection and reducing unnecessary patient recalls as demonstrated by the large clinical study conducted by Getty *et al.* [31]. In that study conducted by Getty *et al.*, 1458 women were enrolled and underwent both standard mammography and stereo mammography examinations [31]. These examinations were then independently interpreted by 5 radiologists. Getty *et al.* found that compared to standard mammography, stereo mammography resulted in a remarkable 23 % increase in true positive detections, and an equally remarkable 46 % reduction in false positive detections [31].

4.3 Computational Stereo for Mammography

The advent of stereo mammographic imaging, while still nascent, has opened the door for the development of computational tools for visualizing and interpreting stereo mammograms. At the crux of computational stereo lies the stereo correspondence or matching problem. Stereo matching is the process of matching each pixel in one of the stereo views to candidate pixels in the other view. Finding the best match yields a positional difference known as disparity arising from the slightly different geometries of the captured views. Once the disparity is estimated at a sufficient number of image locations in the reference view, it is possible to reconstruct the geometry of the original 3-D scene using knowledge of the imaging geometry. As discussed in Chapter 1, the problem of reliably estimating a dense disparity map from a pair of stereo mammogram images is very important since this is the first step towards developing computational tools for interpreting stereo mammograms.

However, estimating the optimal disparity between the two views of a stereo pair is an ill-posed problem. A pixel in the reference view could have multiple candidate matches (often known as the problem of non-uniqueness) in a 2-D search window in the other matching view. The size of this search window is not known *a priori*. To simplify the matching process, it is often assumed that the epipolar constraint holds, that the corresponding image rows are registered. This assumption can be asserted under a non-vergent parallel baseline imaging geometry, but only holds approximately for a vergent imaging geometry provided the angle of separation between the views is small (typically less than 10-15 degrees). However, even with the epipolar assumption in place, the disparity estimation problem is complicated by factors such as non-uniqueness of matches, half occlusions (points of the 3-D scene seen only in one image and not in the other), and transparency (seeing through objects). These kinds of phenomena occur throughout typical stereo mammogram images. We next review some of the relevant work on disparity estimation, mainly for stereo pairs of optical images.

4.4 Relevant Work On Disparity Estimation

A detailed review of disparity estimation is provided in [96] and [17]. Broadly speaking, existing disparity estimation algorithms are founded on the premise of brightness (or color) constancy, i.e., the projected brightness (or color) of 3-D scene points visible on both the stereo views are similar. The brightness constancy assumption is often handled via a matching/similarity cost function that assigns a cost based on the projected brightness values at the corresponding points of the two views. It should be noted that other matching cost functions such as those based on luminance derivatives and mutual information have been explored to handle images having radiometric differences resulting in unequal projected brightness values [41]. However, brightness constancy alone is insufficient to reliably estimate the disparity due to problems associated with non-uniqueness and half occlusions as described before. As a result, the brightness constancy assumption is often augmented with a disparity smoothness assumption, which penalizes large changes in the disparity gradient. The disparity smoothness assumption is based on the premise that natural 3-D scenes are composed of objects and surfaces that are largely piecewise smooth. With these assumptions in place, stereo algorithms fall under two broad categories: 1) local algorithms, and 2) global algorithms.

Local algorithms are typically window-based, where the best matching disparity at a given pixel is dependent on the intensity (or color) values within a local window [96]. The disparity smoothness assumption is typically enforced by aggregating support from neighboring pixels within the local window. Global algorithms pose the disparity estimation problem as an optimization problem [96], where in an energy functional containing a brightness constancy term, or photometric constraint, and the disparity smoothness term is minimized using an optimization algorithm such as simulated annealing [6], graph cuts [15], or variational level sets [8].

Disparity estimation algorithms are sometimes also categorized as sparse or dense algorithms [96]. Sparse algorithms typically employ a feature detector such as an edge or a curvilinear structure detector on the reference image and estimate disparity only at points of high confidence feature extraction. The motivation behind sparse algorithms is to eliminate matching in textureless regions where disparity estimation is often compounded by non-unique matches. Dense stereo algorithms attempt to estimate disparity at every point of the reference image. Disparity estimation algorithms could also be either single scale or multi-scale. Single-scale algorithms operate only on the original image resolution, while multi-scale algorithms (e.g., [8], [6]) employ a hierarchical coarse-to-fine approach, where the disparity estimated at a coarser scale is supplied as an initial estimate to the disparity estimation process at the subsequent finer scale. Low-pass and band-pass image pyramids [1] are commonly employed in multi-scale stereo algorithms. The Middlebury website `http://vision.middlebury.edu/stereo/` contains a comprehensive evaluation of various kinds of stereo algorithms on benchmark natural scene data sets.

Computational stereo and disparity estimation has also been studied for applications involving microscopic and medical images. For instance, Kim, Bovik, and Aggarwal [49] presented one of the earlier algorithms concerning the estimation of disparity in stereo images of biological specimens obtained from a stereo light microscope. Their's was a sparse algorithm that estimated disparity at points of high gradient magnitude and subsequently recovered the 3-D shapes of contours evident on vascular cast stereo images. One of the first algorithms for stereo mammography was proposed by Chelberg *et al.* [23]. They used Laplacian of Gaussian (LOG) filters for performing stereo matching on a stereo mammography pair of a digital phantom. Corresponding points were matched only between zero-crossings of the same sign and roughly the same orientation in the left and the right images, thereby resulting in sparse disparity estimates. Other research groups have proposed algorithms for extracting corresponding features from two-view stereotactic biopsy mammograms and standard two-view (mediolateral oblique and craniocaudal) mammograms for use in computer-aided diagnosis of breast cancer (e.g., [102] and [42]). However, this is a much harder problem since the two views are acquired over a very wide angle (between 30-60 degrees) with different amounts of compression.

4.5 Disparity Estimation for Stereo Mammography

A peculiarity of mammograms is the presence of many singularities in the form of curvilinear structures of various lengths, widths, and tortuosities, which exhibit a complex occlusion pattern. This is unlike what is seen in optical images of typical natural scenes, which are largely comprised of piece-wise smooth surfaces. As discussed in the previous section, computational stereo algorithms for natural scenes exploit this property and constrain the resulting disparity to be piece-wise smooth. However, in the case of stereo mammograms, piece-wise smoothness is violated at singular points often arising out of a network of curvilinear structures comprised of vasculature, ducts, lobules and spicules. For example, Fig. 4.1 illustrates a pair of stereo mammogram im-

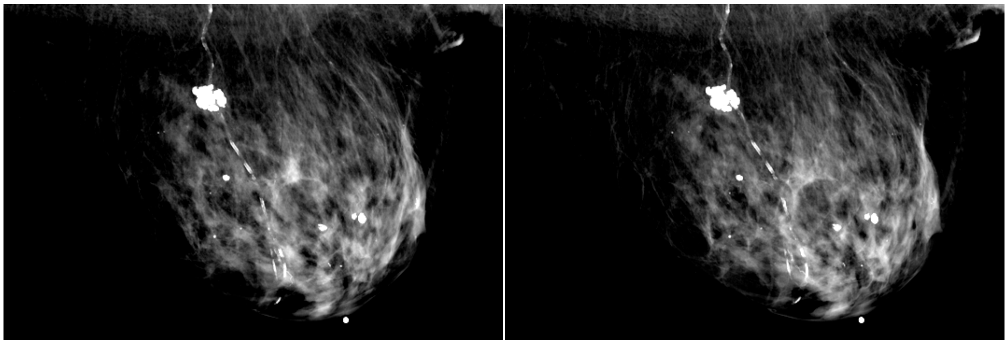

Figure 4.1: A pair of stereo mammogram images.

ages. Stereo acute observers can fuse this pair without using any stereoscopic devices by just crossing their eyes. The resulting 3-D view in the brain, also known as the cyclopean view, is comprised of a number of curvilinear structures lying in different depth planes. Preserving these curvilinear structures in the disparity space is important in order to estimate the depth at which these structures lie. Indeed, the reliable detection of curvilinear structures in mammograms has been a widely studied problem and continues to be of interest for developing robust computer-aided detection (CADe) and diagnosis (CADx) algorithms (e.g., [111] and [68]). Towards solving this difficult aspect of the problem,we have developed a new stereo model that employs a novel singularity index to better estimate disparity along singular points arising out of isolated impulses and curvilinear structures [71], [72], [73]. The singularity index can be configured to detect structures such as impulse masses in a 1-D signal or curvilinear masses in images, while rejecting step edges. It can also be configured to do the opposite. The following chapter discusses the singularity index in detail.

Chapter 5

A New Singularity Index

5.1 Introduction

Detection of singularities in images is a widely studied problem in computer vision and image analysis, since singularities correspond to luminance discontinuities and provide evidence of object contours and surface boundaries. Two types of singularities are often encountered: 1) edge (jump) singularities, and 2) impulse singularities such as curvilinear ridges in images. Impulses or curve-like singularities often arise from curvilinear objects that exist at fine scales. Locating curve-like singularities is important in many applications. For example in mammography, reliable detection of impulse singularities arising from curvilinear structures such as blood vessels, ducts, and spicules is important for CADe applications. Other applications that rely on the detection of impulse singularities include analysis of filaments in images of biological specimens and astronomical bodies, and roads and river deltas in satellite images [54].

Many approaches have been proposed for detecting and localizing singularities in images. The history of edge detection dating from Roberts [89] was greatly advanced by Marr and Hildreth [67] and Canny [20] who employed smoothed Laplacian and gradient operators to detect jump discontinuities. Mallat and Hwang [64] studied singularity detection in the context of wavelet theory. They characterized the Lipschitz regularity of the wavelet transform modulus extrema across scales and showed that the Lipschitz exponent could be used to reveal whether the signal varied smoothly, or whether there was an edge, or an impulse like singularity [64]. Lindeberg [60], [61] and Steger [100] presented a general scale-space framework for detecting edges and ridges.

In this chapter, we present a new ratio index for the detection of impulse singularities in signals of arbitrary dimensionality [71], [73]. The singularity index is motivated by the need to better estimate disparity along curvilinear structures evident on a pair of stereo mammogram images. The index is inspired by conditions put forth by Lindeberg [60], [61] and Steger [100], by Canny's approach to edge detection [20], and by an energy operator developed by Teager and Kaiser [44]. The index employs steerable directional derivatives of gaussians yielding a simple and computationally efficient multi-scale framework. Moreover, the singularity index is scale-invariant. This property is very useful for detecting curvilinear structures existing at different scales. We next describe the proposed index in 1-D.

5.2 Proposed Singularity Index in 1-D

We first define the new singularity index in 1-D and in multiple dimensions afterwards. Let $f(x), x \in R$, and $f'(x)$ and $f''(x)$ denote first and second order derivatives, respectively. Then define the dimensionless ratio index

$$(\psi f)(x) = \frac{|f(x)f''(x)|}{C + |f'(x)|^2} \tag{5.1}$$

where $C \in R$. The index ψ responds strongly to impulse singularities, where the twice-derivative is large, but weakly to step singularities, where the once-derivative is large. Where the once-derivative is small, the denominator has little effect, suggesting the nominal value $C = 1$ although other criteria such as noise might promote other choices. For simplicity, assume $(\psi f)(x) = \frac{|f(x)f''(x)|}{1+|f'(x)|^2}$.

To ensure invariance to local DC offset, the function $f(x)$ is locally debiased prior to computing $(\psi f)(x)$. This is done by everywhere subtracting the local mean computed using a large, unit area gaussian filter. The scale λ of this gaussian may be chosen as follows: consider a smoothed impulse of height K and scale w: $f(x) = Ke^{\frac{-x^2}{2w^2}}$ (Fig. 5.1, left). The locally debiased signal is $\hat{f}(x) = f(x) - g_\lambda(x) * f(x) = Ke^{\frac{-x^2}{2w^2}} - \frac{Kw}{\sqrt{\lambda^2+w^2}}e^{\frac{-x^2}{2(\lambda^2+w^2)}}$. At $x = 0$, $f(0) = K$, and $\hat{f}(0) = K - \frac{Kw}{\sqrt{\lambda^2+w^2}}$. By choosing λ so that $|f(0) - \hat{f}(0)| \leq \epsilon K$, where $\epsilon \in (0, 1]$, the lower bound $\lambda \geq w\sqrt{\frac{1-\epsilon^2}{\epsilon^2}}$ is arrived at. Reasonable values of $w = 1.5$ and $\epsilon = 0.2$ yield $\lambda \geq 7.34$. We next analyze the behavior of the index on 1-D impulse and edge profiles.

5.2.1 1-D Impulse Profile

Model a smoothed 1-D impulse as before by a gaussian of height $K > 0$ and scale w (see Fig. 5.1, left): $f(x) = Ke^{\frac{-x^2}{2w^2}}$. Then the index evaluates to

$$(\psi f)(x) = \frac{\left|\frac{K^2}{w^2}\right| \left|\frac{x^2}{w^2} - 1\right| e^{\frac{-x^2}{w^2}}}{1 + \left|\frac{K^2}{w^4}\right| x^2 e^{\frac{-x^2}{w^2}}}. \tag{5.2}$$

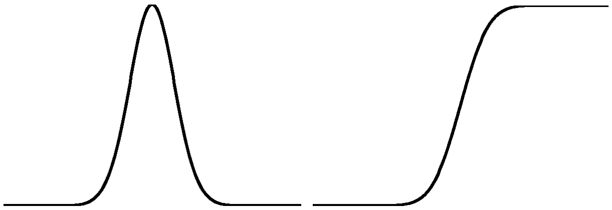

Figure 5.1: 1-D impulse (left) and edge (right) profiles.

At $x = 0$, $(\psi f)(0) = \left|\frac{K^2}{w^2}\right|$. As $K \uparrow$, or $w \downarrow$, $(\psi f)(0) \uparrow$. The singularity index favors sharp impulses. As $w \to \infty$ or $x \to \infty$, then $\psi \to 0$. In the absence of other stimuli, the index vanishes with increased smoothness of, or distance from the impulse. It responds to both positive going and negative going impulses, although polarity is easily retained.

5.2.2 1-D Edge Profile

Model a 1-D edge profile as a step $u(x)$ of height $K > 0$ smoothed by a gaussian $g_w(x)$ (see Fig. 5.1, right):

$$f(x) = Ku(x) * g_w(x) = K\Phi\left(\frac{x}{w}\right), \tag{5.3}$$

where $\Phi\left(\frac{x}{w}\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{x}{w}} e^{\frac{-t^2}{2}}\, dt$. In this case, the singularity index evaluates to

$$(\psi f)(x) = \frac{\left(\left|\frac{K^2}{2\pi w^4}\right| |x| \left(\int_{-\infty}^{x} e^{\frac{-t^2}{2w^2}}\, \mathrm{d}t\right) \left(e^{\frac{-x^2}{2w^2}}\right)\right)}{1 + \left(\left|\frac{K}{\sqrt{2\pi} w}\right|^2 e^{\frac{-x^2}{w^2}}\right)}. \tag{5.4}$$

At $x = 0$, the index vanishes: $(\psi f)(0) = 0$, and near the edge the index decreases with K.

5.2.3 Side-lobe Response

Lemma 1: For an impulse whose cross-section profile is a gaussian, the peak side-lobe response $(\psi f)_s$ of the singularity index is bounded by 1.

Proof: First notice that in (5.2), $(\psi f)(w) = 0$. The side-lobes occur at $|x| > w$. Let $A = \frac{K^2}{w^2}$ and $y = \frac{x^2}{w^2}$. Substituting these in (5.2) yields: $(\psi f)(y) = \frac{A|y-1|e^{-y}}{1+Aye^{-y}}$.

Case 1: $y - 1 > 0 \Rightarrow (\psi f)(y) = \frac{A(y-1)e^{-y}}{1+Aye^{-y}}$

Differentiating ψ w.r.t y and equating to 0 yields $(y-2)e^y = A$, the solution of which is the location of the peak side-lobe response. Substituting for A in ψ yields the peak side-lobe response $(\psi f)_s = \frac{y-2}{y-1} = \frac{x^2-2w^2}{x^2-w^2} \leq 1 \ \forall \ |x| > w$.

Case 2: $y - 1 < 0 \Rightarrow (\psi f)(y) = \frac{A(1-y)e^{-y}}{1+Aye^{-y}}$

Differentiating ψ w.r.t y and equating to 0 again yields $(y-2)e^y = A$. Following similar (reversed) reasoning we find $|(\psi f)_s| \leq 1$.

From *Lemma 1*, $(\psi f)_s$ is bounded by 1 and does not depend on the impulse height K. By comparison, the simple impulse detector $f''(x)$ has a peak side-lobe response $\frac{0.446K}{w^2}$ at $x = \pm\sqrt{3}w$, which increases linearly with K for a fixed scale.

The singularity index does produce a small, but undesirable side-lobe response to edges. Rewrite (5.4) as $(\psi f)(z) = \frac{\left|\frac{K^2}{w^2}\Phi(z)\phi'(z)\right|}{1+\frac{K^2}{w^2}\phi^2(z)}$, where $z = \frac{x}{w}$, $\phi(z) = \frac{1}{\sqrt{2\pi}}e^{\frac{-z^2}{2}}$ and $\Phi(z) = \int_{-\infty}^{z} \phi(t)\, dt$ are the standard normal probability density and cumulative distribution functions, respectively. Differentiating $(\psi f)(z)$ w.r.t z and equating to 0 yields a substitution for $\Phi(z)$, which when used in $(\psi f)(z)$ yields the peak edge side-lobe response $(\psi f)_{Es} = \frac{K^2z^2\phi^2(z)}{K^2(z^2+1)\phi^2(z)-w^2(z^2-1)}$.

Unlike the impulse side-lobe response, the edge side-lobe response is not bounded by 1 and depends on K for a fixed scale w. However, the peak edge side-lobe response can be substantially attenuated as described next.

For some constant $a \in \mathrm{R}$, compute the first derivative response using a scale aw so that the singularity index becomes

$$(\psi f)(x) = \frac{|f(x)f''(x)|}{1 + |f'_a(x)|^2}. \tag{5.5}$$

To choose a, consider the edge model (5.3), where now $f'_a(x) = \frac{K}{aw}\phi(\frac{x}{aw}) = \frac{K}{\sqrt{2\pi}aw}e^{\frac{-x^2}{2a^2w^2}}$, and define $P(a) = aw(f'_a(x))^2$. The optimal value for a is then taken to be $a^* = \underset{\mathrm{a}}{\operatorname{argmax}}\ P(a) = \frac{\sqrt{2}x_0}{w}$, where x_0 is the location where the numerator $|ff''|$ is maximized in response to the step edge. Numerical solution yields $x_0 = 1.2554w$ and $a^* = 1.7754$. Extensive simulations using $a = 1.7754$ suggest that the peak side-lobe edge response of the singularity index computed with $f'_a(x)$ in (5.5) is upper bounded by 3.7. By comparison, $f''(x)$ has a peak side-lobe edge response $\frac{K}{\sqrt{2\pi e}w^2}$ at $x = \pm w$, which increases linearly with K for a fixed scale w.

5.2.4 Generalized Singularity Index

It is possible to control the scale of the singularity index by smoothing the signal first:

$$(\psi f)(x, \sigma) = \frac{|g_\sigma * f(x)||g''_\sigma * f(x)|}{1 + |g'_{a\sigma} * f(x)|^2}, \tag{5.6}$$

where, g is a smoothing filter such as an unit-area gaussian: $g_\sigma(x) = \frac{1}{\sqrt{2\pi}\sigma}e^{\frac{-x^2}{2\sigma^2}}$. The first derivative of the smoothed signal is computed at a scale $a\sigma$ to attenuate the edge side-lobe responses. The use of gaussian filters stabilizes derivative computations and reduces noise. The concept of the index (5.6) can be generalized to detect other types of discontinuities such as edges. The k^{th}-order index

$$(\psi^k f)(x,\sigma) = \frac{|g_\sigma^{k-1} * f(x)||g_\sigma^{k+1} * f(x)|}{1 + |g_{a\sigma}^k * f(x)|^2} \tag{5.7}$$

includes (5.6) as a special case ($k = 1$). Odd integer values of k yield impulse detectors, while even integer values of k yield edge detectors.

5.2.5 Multi scale analysis

The smoothed singularity index (5.6) is easily extended to detect impulses at multiple scales. Consider a smoothed gaussian impulse of height $K > 0$ and scale w, i.e. $f(x) = Ke^{\frac{-x^2}{2w^2}}$. Then $f(x) * g_\sigma(x)$ is also gaussian: $f(x) * g_\sigma(x) = \frac{Kw}{\sqrt{\sigma^2+w^2}} e^{\frac{-x^2}{2(\sigma^2+w^2)}}$. Similarly, it can be shown that the first derivative response is $(f(x) * g_{a\sigma}(x))' = \frac{-Kwx}{((a\sigma)^2+w^2)^{3/2}} e^{\frac{-x^2}{2((a\sigma)^2+w^2)}}$ and the second derivative response is $(f(x) * g_\sigma(x))'' = \frac{-Kw}{(\sigma^2+w^2)^{3/2}} \left(\frac{-x^2 e^{\frac{-x^2}{2(\sigma^2+w^2)}}}{(w^2+\sigma^2)} + e^{\frac{-x^2}{2(\sigma^2+w^2)}} \right)$. Then $(\psi f)(0,\sigma) = \left| \frac{-K^2w^2}{(\sigma^2+w^2)^2} \right| = \frac{K^2w^2}{(\sigma^2+w^2)^2}$.

Now consider the scale normalized index $(\psi_{norm} f)(x,\sigma) = \sigma^{2\gamma}(\psi f)(x,\sigma)$, where $\gamma \in R$. The scale-normalized index is motivated by the scale-space selection methodology based on local extrema over scales of γ-normalized derivatives that was proposed by Lindeberg in [61]. Then $(\psi_{norm} f)(0,\sigma)$ is maximized when $\sigma = \sqrt{\frac{2\gamma}{4-2\gamma}}w$, which matches the smoothed impulse width when $\gamma = 1$. Hence, the scale normalized index $(\psi_{norm} f)(x,\sigma) = \sigma^2(\psi f)(x,\sigma)$

attains a scale-space maxima at $\sigma = w$. Substituting this into the scale-normalized index yields $(\psi_{norm} f)(0, w) = \frac{K^2}{4}$, a constant independent of w. Thus, the scale-normalized singularity index is scale-invariant.

It is worth comparing the scale-normalized index to Lindeberg's γ-normalized *maximum eigenvalue of the Hessian* ridge strength measure, $M_{\gamma-norm}L$, defined in equation (46) in [60] for a smoothed gaussian impulse $f(x) = Ke^{\frac{-x^2}{2w^2}}$. As shown in [60], in order that the maximizing scale equal the width of the impulse, $\gamma = 3/4$. Substituting this into $M_{\gamma-norm}L$ yields a response of $\frac{K}{2^{3/2}\sqrt{w}}$ at the origin, which is not scale invariant.

5.3 Proposed Singularity Index in 2-D

The singularity index (5.1) is easily extended to higher dimensions. Let $f(\mathbf{x}), R^n \to R$ be the n-D signal, ∇^2 be the n-D Laplacian $\nabla^2 f(\mathbf{x}) = \sum_{i=1}^{n} \left(\frac{\partial^2 f(\mathbf{x})}{\partial x_i^2} \right)$ and ∇ be the n-D gradient $\nabla f(\mathbf{x}) = \left(\frac{\partial f(\mathbf{x})}{\partial x_1}, \ldots.., \frac{\partial f(\mathbf{x})}{\partial x_n} \right)^T$. Then define the n-D singularity index as

$$(\psi f)(\mathbf{x}) = \frac{|f(\mathbf{x}) \nabla^2 f(\mathbf{x})|}{1 + |\nabla f(\mathbf{x})|^2} \tag{5.8}$$

assuming the nominal value $C = 1$. In [60], [100], ridge detection in 2-D images used the predicate that the first derivative in the direction of the maximum eigen value of the Hessian matrix of the signal be 0, while the second derivative be non zero. The 2-D singularity index ψ naturally embodies these conditions tying the first and second derivatives together in a simple and elegant way to yield a dimensionless quantity. The polarity can easily be retained by the slight modification: $(\psi f)(\mathbf{x}) = \frac{|f(\mathbf{x})|\nabla^2 f(\mathbf{x})}{1+|\nabla f(\mathbf{x})|^2}$, which is equivalent to multiplying (8) by $sgn(\nabla^2 f(\mathbf{x}))$. We next consider the behavior of the 2-D singularity index when applied to a 2-D smoothed isotropic impulse and and a 2-D line impulse profile.

5.3.1 2-D Isotropic Impulse

Consider the detection of a simple 2-D singularity ($n = 2$) first. As in the 1-D case, model a smoothed 2-D impulse by an isotropic gaussian of strength $K > 0$: $f(x, y) = Ke^{\frac{-(x^2+y^2)}{2w^2}}$.The 2-D singularity index evaluates to

$$(\psi f)(x, y) = \frac{\left|\frac{K^2}{w^2}\right| \left|(\frac{x^2+y^2}{w^2} - 2)\right| e^{\frac{-(x^2+y^2)}{w^2}}}{1 + \left(\frac{K^2}{w^4} e^{\frac{-(x^2+y^2)}{w^2}} (x^2 + y^2)\right)}. \tag{5.9}$$

At $x = y = 0$,

$$(\psi f)(0, 0) = 2\left|\frac{K^2}{w^2}\right|. \tag{5.10}$$

Again, as in the 1-D case, as $K \uparrow$, or $w \downarrow$, $(\psi f)(0, 0) \uparrow$.

5.3.2 2-D Line Impulse Profile

Model a smoothed 2-D line impulse profile by a 2-D isotropic gaussian of strength $K > 0$ and scale w convolved with the impulse sheet $\delta(x)$: $f(x, y) = \frac{K}{\sqrt{2\pi}w} e^{\frac{-(x^2+y^2)}{2w^2}} * \delta(x) = Ke^{\frac{-x^2}{2w^2}}, \forall y$. Thus, conclusions regarding the response of the singularity index to the 2-D line impulse profile are the same as for the 1-D impulse profile. Generally, the new singularity index responds strongly to impulses along line and smooth curvilinear masses.

5.3.3 Practical Implementation

Direct application of (5.8) to real discrete space images is not practical owing to instabilities in the derivative computations. However, smoothing the image first, then computing partial derivatives (or combining derivatives with the smoothing filter) produces the scalable index $(\psi_\sigma f)(x,y) = \frac{|g_\sigma(x,y)*f(x,y)||\nabla^2 g_\sigma(x,y)*f(x,y)|}{1+|\nabla g_{a\sigma}(x,y)*f(x,y)|^2}$, where the gradient of the smoothed signal is again computed at a scale $a\sigma$. The sensitivity of the index can be further improved by adopting a design mechanism inspired by Canny for edge detection [20]. First, determine the direction θ at each pixel along which the second derivative of the gaussian filtered image attains a local extrema, which is a good estimate of the direction orthogonal to an curvilinear singularity. Once this direction is estimated, evaluate the responses of the gaussian derivative filters along this direction and compute the ratio index as follows:

$$(\psi f)(x,y,\sigma) = \frac{|f_{0,\theta,\sigma}(x,y) f_{2,\theta,\sigma}(x,y)|}{1+|f_{1,\theta,a\sigma}(x,y)|^2} \tag{5.11}$$

In (5.11), $f_{0,\theta,\sigma}(x,y)$, $f_{1,\theta,a\sigma}(x,y)$, and $f_{2,\theta,\sigma}(x,y)$ are the responses to the zero, first and second order gaussian derivative filters along the direction specified by $\theta(x,y)$ and at scale σ. To estimate $\theta(x,y)$, deploy an isotropic gaussian filter and exploit the steerable property of gaussian directional derivatives as described next.

The response of a second order directional derivative of an isotropic gaussian can be synthesized as a linear combination of the responses of 3 second order derivatives along evenly spaced directions [29]:

$$\begin{aligned} f_{2,\theta,\sigma}(x,y) = f * G_{2,\theta,\sigma} = 1/3(1 + 2\cos(2\theta))(J_{2,\theta=0,\sigma}) + \\ 1/3(1 - \cos(2\theta) + \sqrt{3}\sin(2\theta))(J_{2,\theta=\pi/3,\sigma}) + \\ 1/3(1 - \cos(2\theta) - \sqrt{3}\sin(2\theta))(J_{2,\theta=2\pi/3,\sigma}) \end{aligned} \qquad (5.12)$$

In (5.12), $G_{2,\theta,\sigma}$ is the derivative of the isotropic gaussian along direction θ at a fixed scale σ, $\theta = 0, \pi/3, 2\pi/3$ are the basis directions, and $J_{2,\theta,\sigma}$ are the responses of the basis filters. Estimating θ along which the squared second order directional derivative attains a local maximum or minimum is straightforward as shown in [29]. Once this estimate of θ is obtained, (5.11) can be applied to compute the singularity index: $f_{2,\theta,\sigma}(x,y)$ is computed according to (5.12), and $f_{1,\theta,a\sigma}(x,y)$ is computed using the steerability of the first order isotropic gaussian derivative (at scale $a\sigma$), which is trivial, and $f_{0,\theta,\sigma}(x,y)$ is just $f(x,y)$ smoothed by the isotropic gaussian function at the given scale σ.

5.3.4 Multi-scale Realization

As in the 1-D case, the use of gaussian derivatives allows for an easy extension of the singularity index (5.11) to detect curvilinear masses at multiple scales. We employ a discrete, coarse-to-fine approach where (5.11) is computed in the order of decreasing scale σ. Automatic scale selection is achieved by selecting the scale yielding the largest scale normalized index value.

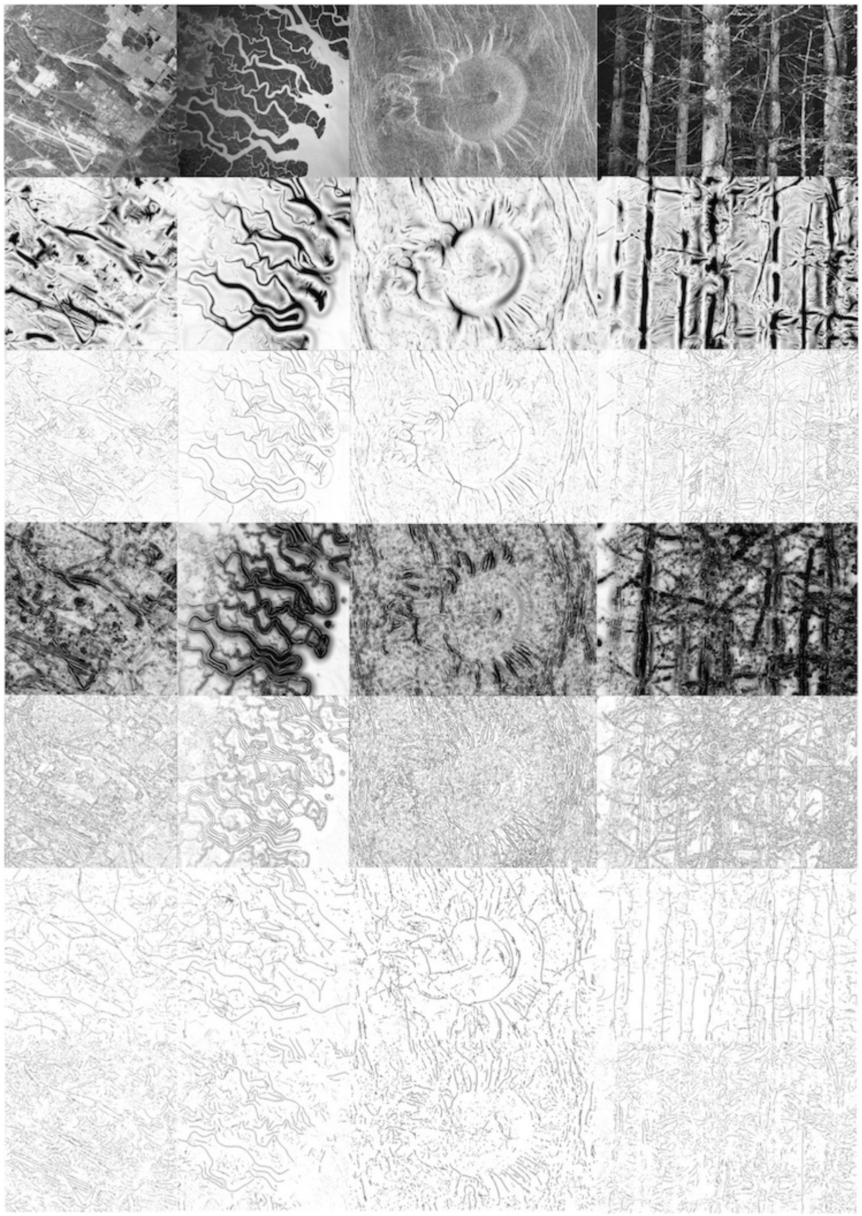

Figure 5.2: Row 1: Original images; row 2: normalized singularity index result; row 3: Non-maxima suppression (NMS) on singularity index result; row 4: normalized second derivative index; row 5: NMS on second derivative index; row 6: Lindeberg's $A_{\gamma-norm}L$ ridge strength measure using 50 scales; and row 7: Lindeberg's $A_{\gamma-norm}L$ ridge strength measure using 6 scales.

5.4 Experiments and Results

In all our experiments, we computed the scale-normalized singularity index i.e. $\sigma^2(\psi f)(x, y, \sigma)$. We fixed the lowest scale σ of the isotropic gaussian to 1.5 pixels, and the constant $a = 1.7754$. Each subsequent coarser scale was larger than the previous finer scale by a factor of $\sqrt{2}$. The number of scales was set to 6. Prior to computing the singularity index (5.11), the image was locally mean debiased using a large unit-volume isotropic gaussian with $\lambda \geq w\sqrt{\frac{1-\epsilon^2}{\epsilon^2}}$ pixels, where $\epsilon = 0.2$, and w was the largest scale. For comparison, we also show results using the scale normalized second derivative index i.e. $\sigma^{1.5} f_{2,\theta,\sigma}(x, y)$ and Lindeberg's *square of the* γ*-normalized eigenvalue difference* ridge strength measure ($A_{\gamma-norm}L$ in equation (51) in [60]). For $A_{\gamma-norm}L$, we used an implementation by Kokkinos *et al.* [53]. It has to be noted that $\sigma^{1.5} f_{2,\theta,\sigma}(x, y)$ exactly corresponds to the ridge strength measure $M_{\gamma-norm}L$ defined by Lindeberg in equation (46) in [60]. The scale normalized second derivative index was computed using the same set of scales as the singularity index, while Lindeberg's $A_{\gamma-norm}L$ measure [60] was computed using the same set of scales as the singularity index and also with a more exhaustive 50 scales (provided as the default setting in the implementation [53]). Further, the implementation provided by [53] applies a square-root transformation to $A_{\gamma-norm}L$ to yield a contour saliency measure whose dimensionality is the same as $\sigma^{1.5} f_{2,\theta,\sigma}(x, y)$. No such transformation is applied to the singularity index since it is dimensionless under the assumption that the constant in the denominator has a dimension $intensity^2/length^2$.

Fig. 5.2 illustrates the results of the singularity index in rows 2 and 3, the scale normalized second derivative measure in rows 4 and 5, and Lindeberg's $A_{\gamma-norm}L$ measure (after the square-root transformation) in rows 6 and 7 on four images containing curvilinear structures - an aerial image (courtesy University of Southern California) in column 1, the Ganges river delta (NASA, courtesy of nasaimages.org) acquired as part of the NASA Human Spaceflight collection in column 2, a volcano on venus (NASA, courtesy of apod.nasa.gov) captured by the Magellan spacecraft in column 3, and an image of pine tree trunks in column 4. Each result illustrates the maximum response computed across all scales at every pixel location. The results in rows 3, 5, 6 and 7 are obtained after applying non-maxima suppression (NMS) along the dominant orientation at each pixel after the maximum response across all scales has been computed. For $\sigma^2(\psi f)(x, y, \sigma)$ and $\sigma^{1.5} f_{2,\theta,\sigma}(x, y)$, we use the steerability of gaussian derivatives to estimate the dominant orientation $\theta(x, y)$, while for $A_{\gamma-norm}L$, [53] uses the eigen directions of the Hessian matrix of the brightness function to estimate $\theta(x, y)$.

The results in Fig. 5.2 illustrate a strong response by the proposed singularity index to impulse-like structures, with suppressed response to edges. The normalized second derivative index (rows 4 and 5) produces significant response to edges as noted by Lindeberg too [60]. Lindeberg's $A_{\gamma-norm}L$ ridge strength measure (rows 6 and 7) produces a better response than the normalized second derivative index, but misses out on fine-scale impulses and lacks the contour continuity produced by the new singularity index (rows 2 and 3).

A broader set of results on a total of 18 high-resolution images can be downloaded from: `http://live.ece.utexas.edu/research/SingularityIndex/SingularityIndex.zip`.

Fig. 5.3 offers a detailed look into the response of the singularity index (after NMS) to the pine tree trunks image when the scale σ of the smoothing gaussian is varied. The response illustrated in Fig. 5.3 is from a single-scale computation of the index. It is evident from Fig. 5.3 that the singularity index nicely captures curvilinear structures existing at different scales. At lower scales, (e.g. $\sigma = 3$ in Fig. 5.3), more detail emerges in the response of the index, since fine-scale curvilinear structures are captured. On the other hand, at higher scales (e.g. $\sigma = 12$ in Fig. 5.3), the response is predominantly from large-scale curvilinear structures.

5.5 Summary

We presented a new steerable, multi-scale singularity index for analyzing impulse singularities in images. Our analyses and experiments reveal promising behavior by the index for detecting impulse-like or ridge curvilinear structures in images. The index is scalable and efficient due to the steerable directional derivatives of isotropic gaussians. Most importantly, the index is scale invariant. We next analyze the detection power of the index in the presence of noise, which is important since it reveals how predictable the index is when applied to an impulse that is submerged in noise.

Figure 5.3: Response of the singularity index (after NMS) to the pine tree trunks image when the scale σ of the smoothing gaussian is varied. Top row: original image (left), NMS index response at scale $\sigma = 3$ (right), Bottom row: NMS index response at scale $\sigma = 6$ (left), NMS index response at scale $\sigma = 12$ (right).

Chapter 6

Noise Analysis of the New Singularity Index

6.1 Introduction

In chapter 5, we defined the nominal 1-D singularity index

$$(\psi f)(t) = \frac{|f(t)f''(t)|}{1 + |f'(t)|^2}, \tag{6.1}$$

the smoothed singularity index

$$(\psi_\sigma f)(t) = \frac{|g_\sigma * f(t)||g''_\sigma * f(t)|}{1 + |g'_\sigma * f(t)|^2}, \tag{6.2}$$

and the scale-normalized version

$$(\psi_{\sigma,norm} f)(t) = \sigma^{2\gamma}(\psi_\sigma f)(t), \tag{6.3}$$

where $f(t), t \in R$ is a 1-D signal, g_σ is a smoothing filter such as a unit-area gaussian of scale σ:

$$g_\sigma(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-t^2}{2\sigma^2}}. \tag{6.4}$$

and g'_σ, and g''_σ are the first and second derivatives of the smoothing gaussian.

Real images such as mammograms are invariably noisy. For the index to be useful, it is important to understand its behavior in the presence of noise. Also, the index defined in (6.2) is a non-negative, real valued operator, i.e. $(\psi_\sigma f)(t) \in R^+$. However, many applications require a binary response, indicating the presence or absence of a curvilinear structure. For example, to analyze the curvilinear structures depicted on a mammogram, it is useful to produce a binary map of the curvilinear structures and then study their properties as in [111]. To produce a binary response from the index, careful selection of thresholds is necessary, especially in a noisy environment.

In this chapter, we shall analyze the impulse detection power of the smoothed 1-D singularity index in the presence of noise [72]. These analyses reveal how predictable the index is when applied to an impulse that is submerged in noise. Our analyses also suggest useful thresholds for producing a binary index response map from images afflicted with noise.

The remainder of this chapter is organized as follows: In Section 6.2, we establish the existence of the low-order moments (mean and variance) of the response of ψ_σ to zero-mean, wide sense stationary (WSS) gaussian random noise. In Section 6.3, we conduct a false alarm analysis, finding the probability of false impulse detection in the presence of noise only followed by an analysis of the probability of true impulse detection for an impulse submerged in noise. Further, since the index is designed to be edge suppressing, we also study the probability of incorrectly detecting an edge submerged in noise in Section 6.4. In Section 6.5, we use the results of our analyses as a guide to determine thresholds for extracting curvilinear structures from example 2-D images afflicted with noise. Finally, we summarize the chapter in Section 6.6.

6.2 Low-Order Moments of ψ_σ

We begin by establishing the existence of the low-order moments (mean and the variance) of the smoothed singularity index (6.2) when applied to a random signal $n(t)$. These moments reveal useful insights into the stability and the predictability of the index when applied to signals submerged in noise.

Assume that $n(t)$ is a zero-mean WSS gaussian random process with auto-correlation function

$$R_n(\tau) = q^2\delta(\tau), \tag{6.5}$$

where the input noise variance is $R_n(0) = q^2$. The assumption of wide sense stationarity and gaussianity allows for a tractable statistical analyses of ψ_σ. We will deal with the slightly more complex non-zero mean case when we study the statistical behavior of the index when applied to a deterministic signal immersed in a random zero-mean noise process.

Let $n_\sigma(t) = n(t) * g_\sigma(t)$, $n'_\sigma(t) = (n(t) * g_\sigma(t))'$, and $n''_\sigma(t) = (n(t) * g_\sigma(t))''$, where $g_\sigma(t)$ is a unit-area gaussian smoothing filter as defined in (6.4). All of the processes $n_\sigma(t)$, $n'_\sigma(t)$ and $n''_\sigma(t)$ are also zero-mean WSS gaussian [80]. Moreover, $n'_\sigma(t)$ is statistically independent of both $n_\sigma(t)$ and $n''_\sigma(t)$. Hence, the singularity index

$$(\psi_\sigma n)(t) = \frac{|n_\sigma(t)||n''_\sigma(t)|}{1 + |n'_\sigma(t)|^2} \tag{6.6}$$

is the ratio of two statistically independent random processes.

Denote the numerator random process in (6.6) by X and the denominator random process by Y. From Theorem 2 in [22], the ratio X/Y has all the moments up to order $\leq q$ provided the following condition is satisfied:

$$Pr\{|Y| < T\} \leq AT^{\frac{pq}{p-q}(1+\alpha)}, \tag{6.7}$$

for any $\alpha, A > 0$, integers $p > q$, and $\forall\ 0 < T < \alpha$. Letting $\alpha = 1$, this condition holds since

$$Pr\{|Y| < T\} = Pr\{|1 + |n'_\sigma|^2| < T\} = 0 \leq AT^{\frac{pq}{p-q}(1+\alpha)} \tag{6.8}$$

$\forall\ A > 0,\ p > q,\ 0 < T < 1$, since $1 + |n'_\sigma|^2 \geq 1$. We next derive approximations for the mean and the variance of $(\psi_\sigma n)(t)$.

The auto-correlation function of the process $n_\sigma(t)$ is $R_{n_\sigma}(\tau) = g_\sigma(\tau) * g_\sigma(-\tau) * R_n(\tau)$, since $n(t)$ is WSS [80]. This evaluates to

$$R_{n_\sigma}(\tau) = \frac{q^2}{2\sqrt{\pi}\sigma} e^{-\frac{\tau^2}{4\sigma^2}}. \tag{6.9}$$

Hence, the variance of $n_\sigma(t)$ is

$$var[n_\sigma(t)] = R_{n_\sigma}(0) = \nu_0^2 = \frac{q^2}{2\sqrt{\pi}\sigma}. \tag{6.10}$$

Further, the variances of the derivative processes $n'_\sigma(t)$ and $n''_\sigma(t)$ can be written [80]

$$var[n'_\sigma(t)] = R_{n'_\sigma}(0) = -R^{(2)}_{n_\sigma}(0) = \nu_1^2 = \frac{q^2}{4\sqrt{\pi}\sigma^3}, \tag{6.11}$$

and

$$var[n''_\sigma(t)] = R_{n''_\sigma}(0) = R^{(4)}_{n_\sigma}(0) = \nu_2^2 = \frac{3q^2}{8\sqrt{\pi}\sigma^5}, \tag{6.12}$$

where $R^{(k)}_{n_\sigma}(\tau) = \frac{d^k R_{n_\sigma}(\tau)}{d\tau^k}$.

In general, it is difficult to derive exact expressions for the statistical mean and variance of (6.6). However, it is possible to re-write (6.6) as

$$(\psi_\sigma n)(t) = \left| \frac{n_\sigma(t) n''_\sigma(t)}{1 + |n'_\sigma(t)|^2} \right|, \tag{6.13}$$

since $1 + |n'_\sigma(t)|^2$ is positive. Further, let

$$(\psi_\sigma n)(t) = |(\widehat{\psi_\sigma n})(t)|, \tag{6.14}$$

where $(\widehat{\psi_\sigma n})(t) = \frac{n_\sigma(t) n''_\sigma(t)}{1+|n'_\sigma(t)|^2}$. The approximate mean and variance of the ratio $(\widehat{\psi_\sigma n})(t)$ can be found using the Taylor expansion [80]:

$$E\left[\frac{X}{Y}\right] = \frac{E[X]}{E[Y]} - \frac{cov[X,Y]}{[E[Y]]^2} + \frac{var[Y]E[X]}{[E[Y]]^3}, \tag{6.15}$$

and

$$var\left[\frac{X}{Y}\right] = \frac{var[X]}{[E[Y]]^2} - 2\frac{cov[X,Y]E[X]}{[E[Y]]^3} + \frac{var[Y][E[X]]^2}{[E[Y]]^4}, \tag{6.16}$$

where X and Y are defined as above and $E[\ .\ ]$ denotes expectation.

It can be shown that [14]

$$E[X] = E[n_\sigma(t) n''_\sigma(t)] = R^{(2)}_{n_\sigma}[0] = -\nu_1^2, \tag{6.17}$$

$$E[Y] = E[1 + (n'_\sigma(t))^2] = 1 + R_{n'_\sigma}(0) = 1 - R^{(2)}_{n_\sigma}(0) = 1 + \nu_1^2. \tag{6.18}$$

Using Isserlis's formula for the product moment coefficient of normally distributed random variables [43], the variance of the numerator random process evaluates to

$$\begin{aligned} var[X] = var[n_\sigma(t) n''_\sigma(t)] &= R_{n_\sigma}(0) R^{(4)}_{n_\sigma}(0) + [R^{(2)}_{n_\sigma}(0)]^2 \\ &= \nu_0^2 \nu_2^2 + \nu_1^4. \end{aligned} \tag{6.19}$$

The variance of the denominator random process is $var[Y] = var[1 + |n'_\sigma(t)|^2] = var[[n'_\sigma(t)]^2]$. Let $m = [n'_\sigma(t)]^2$. Then,

$$var[[n'_\sigma(t)]^2] = var[m] = E[m^2] - [E[m]]^2. \tag{6.20}$$

Since $n'_\sigma(t)$ is zero-mean WSS gaussian, it can be shown that [80]

$$E[m^2] = R_m(0) = 3[R_{n'_\sigma}(0)]^2 = 3[R^{(2)}_{n_\sigma}(0)]^2, \tag{6.21}$$

and

$$[E[m]]^2 = [E[n'_\sigma(t)]^2]^2 = [R_{n'_\sigma}(0)]^2 = [R^{(2)}_{n_\sigma}(0)]^2. \tag{6.22}$$

Hence,

$$\begin{aligned} var[Y] &= var[1 + |n'_\sigma(t)|^2] = var[m] \\ &= 3[R^{(2)}_{n_\sigma}(0)]^2 - [R^{(2)}_{n_\sigma}(0)]^2 \\ &= 2[R^{(2)}_{n_\sigma}(0)]^2 = 2\nu_1^4. \end{aligned} \tag{6.23}$$

Substituting (6.17), (6.18), (6.19), (6.23), and $cov[X, Y] = 0$ (since X and Y are statistically independent) in (6.15) and (6.16) yields

$$E[(\hat{\psi_\sigma} n)(t)] = \frac{-\nu_1^2}{1 + \nu_1^2} + 2\left(\frac{-\nu_1^2}{1 + \nu_1^2}\right)^3. \tag{6.24}$$

and

$$var[(\hat{\psi_\sigma} n)(t)] = \frac{\nu_0^2\nu_2^2 + \nu_1^4}{(1 + \nu_1^2)^2} + \frac{2\nu_1^8}{(1 + \nu 1^2)^4}. \tag{6.25}$$

The mean and variance of $(\hat{\psi_\sigma} n)(t)$ place bounds on the moments of the singularity index $(\psi_\sigma n)(t)$ as a consequence of Jensen's inequality:

$$\begin{aligned} E[(\psi_\sigma n)(t)] &> E[(\hat{\psi_\sigma} n)(t)] \\ var[(\psi_\sigma n)(t)] &< var[(\hat{\psi_\sigma} n)(t)]. \end{aligned} \tag{6.26}$$

Finally, Fig. 6.1 illustrates the mean (top) and variance (bottom) of $(\hat{\psi}_\sigma n)(t)$ computed using the Taylor approximation for different values of the input noise variance q^2 as a function of the smoothing filter scale σ. To ascertain the veracity of the Taylor approximation, we plot the mean and variance of $(\hat{\psi}_\sigma n)(t)$ computed empirically for simulated zero-mean white gaussian noise at the same values of q^2 and σ. As evident from Fig. 6.1, the mean and variance of $(\hat{\psi}_\sigma n)(t)$ estimated using the Taylor approximation and empirical simulation are nearly identical. It is also evident from Fig. 6.1 that for each input noise variance, the mean of $(\hat{\psi}_\sigma n)(t)$ approaches zero as the smoothing filter scale increases, while the variance of $(\hat{\psi}_\sigma n)(t)$ decreases as the smoothing filter scale increases. These properties are desirable since it suggests that as the noise is reduced by increased smoothing, the expected value and the variance of the index response tend towards zero.

6.3 False Alarm and True Impulse Detection Probabilities

6.3.1 False Alarm Probability

Here, we find the probability of false impulse detection given a threshold, when the scale-normalized smoothed singularity index (6.3) ($\gamma = 1$) is applied to a zero-mean WSS gaussian random process $n(t)$. This is a simple test of whether an impulse singularity will be detected when none is present (false alarm). We only deal with the scale-normalized case, since in real applications the scale-normalized index is most useful for detecting impulses and

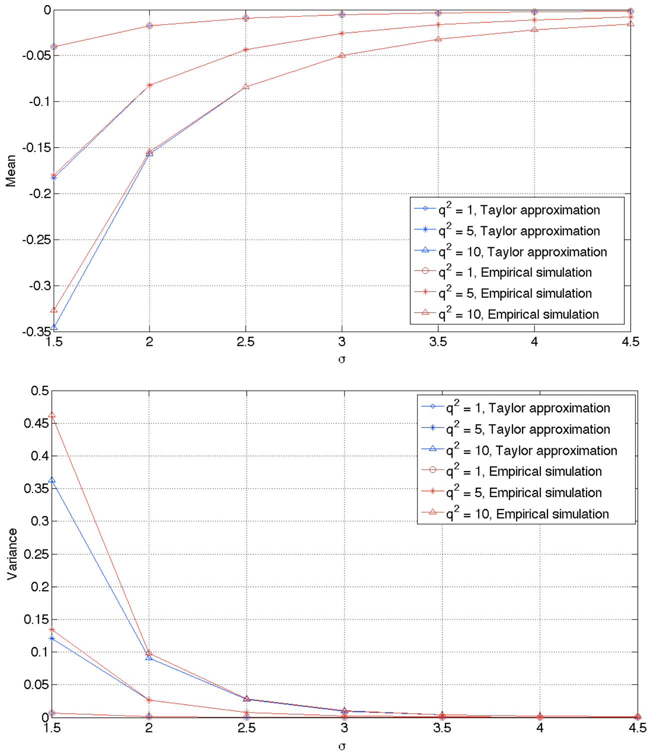

Figure 6.1: Mean (top) and variance (bottom) of $\hat{\psi}_\sigma[n(t)]$ estimated using the Taylor approximation (blue) and empirical simulations (red) for different values of the input noise variance q^2 as a function of the smoothing filter scale σ.

for obtaining an estimate of their scale/width. Lindeberg [61] first presented scale-normalization in the context of ridge detection and stated that it was necessary to normalize the derivatives computed at each scale of the smoothing gaussian in order to compare the derivative magnitudes across scales and to obtain an estimate of the width of the ridge. We begin by first finding the false alarm probability for the smoothed singularity index (6.2) and subsequently use this result to find the false alarm probability for the scale-normalized index (6.3).

As discussed earlier, the random process $(\psi_\sigma n)$ (6.6) is the ratio of two statistically independent random processes $X = |n_\sigma||n''_\sigma|$ and $Y = 1 + |n'_\sigma|^2$, i.e. $(\psi_\sigma n) = X/Y$. Since n'_σ is a zero-mean gaussian random process, the probability density function (PDF) of Y is given by [80]

$$f_Y(y) = \frac{1}{\sqrt{2\pi(y-1)}\nu_1} exp\left(-\frac{(y-1)}{2\nu_1^2}\right), \ y > 1. \tag{6.27}$$

The numerator random process is of the form $X = |\hat{X}|$, where $\hat{X} = n_\sigma n''_\sigma$. Since both n_σ and n''_σ are correlated zero-mean gaussian random processes with correlation coefficient $\rho = \frac{-\nu_1^2}{\nu_0\nu_2}$, their product $\hat{X}$ has a PDF that is given by [14]

$$f_{\hat{X}}(x) = \frac{1}{\pi\sqrt{\nu_0^2\nu_2^2 - \nu_1^4}} exp\left(\frac{-x\nu_1^2}{\nu_0^2\nu_2^2 - \nu_1^4}\right) K_0\left(\frac{|x|\nu_0\nu_2}{\nu_0^2\nu_2^2 - \nu_1^4}\right), \tag{6.28}$$

where $K_0(\ . \ )$ is the modified Bessel function of the second kind and order zero. The PDF of $X = |\hat{X}|$ is given by

$$f_X(x) = f_{\hat{X}}(x) + f_{\hat{X}}(-x), \ x > 0. \tag{6.29}$$

Hence,

$$\begin{aligned} f_X(x) = \frac{1}{\pi\sqrt{\nu_0^2\nu_2^2 - \nu_1^4}} K_0\left(\frac{|x|\nu_0\nu_2}{\nu_0^2\nu_2^2 - \nu_1^4}\right) \\ .\left\{exp\left(\frac{-x\nu_1^2}{\nu_0^2\nu_2^2 - \nu_1^4}\right) + exp\left(\frac{x\nu_1^2}{\nu_0^2\nu_2^2 - \nu_1^4}\right)\right\},\ x > 0. \end{aligned} \tag{6.30}$$

Determining a closed form expression for the PDF of $(\psi_\sigma n)$ is generally difficult. However, numerical evaluation of the probabilities involving $(\psi_\sigma n)$ is feasible. Specifically, evaluate the probability

$$Pr\{(\psi_\sigma n) > \tau\} = Pr\{X/Y > \tau\}, \tag{6.31}$$

where $\tau > 0$ is a threshold. The probability in (6.31) can be determined as [80]

$$\begin{aligned} Pr\{X/Y > \tau\} = \int_{x=0}^{\infty}\int_{y=0}^{\frac{x}{\tau}} f_{XY}(x,y)\ \mathrm{d}y\ \mathrm{d}x\ + \\ \int_{x=-\infty}^{0}\int_{y=\frac{x}{\tau}}^{0} f_{XY}(x,y)\ \mathrm{d}y\ \mathrm{d}x. \end{aligned} \tag{6.32}$$

Since X and Y are statistically independent of each other, (6.32) becomes

$$\begin{aligned} Pr\{X/Y > \tau\} = \int_{x=0}^{\infty}\left(\int_{y=0}^{\frac{x}{\tau}} f_Y(y)\ \mathrm{d}y\right) f_X(x)\ \mathrm{d}x\ + \\ \int_{x=-\infty}^{0}\left(\int_{y=\frac{x}{\tau}}^{0} f_Y(y)\ \mathrm{d}y\right) f_X(x)\ \mathrm{d}x. \end{aligned} \tag{6.33}$$

The second double integral in (6.33) evaluates to 0 since $f_Y(y) = 0\ \forall\ y \leq 1$. Hence, (6.33) can be written

$$Pr\{X/Y > \tau\} = \int_{x=0}^{\infty}\left(\int_{y=1}^{\frac{x}{\tau}} f_Y(y)\ \mathrm{d}y\right) f_X(x)\ \mathrm{d}x. \tag{6.34}$$

Consider the inner integral

$$\int_{y=1}^{\frac{x}{\tau}} f_Y(y)\ \mathrm{d}y = \int_{y=1}^{\frac{x}{\tau}} \frac{1}{\sqrt{2\pi(y-1)}\nu_1} exp\left(-\frac{(y-1)}{2\nu_1^2}\right)\ \mathrm{d}y. \tag{6.35}$$

On introducing the change of variable $z = y - 1$, (6.35) can be written as

$$\int_{y=1}^{\frac{x}{\tau}} f_Y(y)\ \mathrm{d}y = \frac{1}{\sqrt{2\pi}\nu_1} \int_{z=0}^{\frac{x}{\tau}-1} \frac{exp\left(-\frac{z}{2\nu_1^2}\right)}{\sqrt{z}}\ \mathrm{d}z. \tag{6.36}$$

From (3.361) in [34], it can be easily seen that

$$\begin{aligned} \frac{1}{\sqrt{2\pi}\nu_1} \int_{z=0}^{\frac{x}{\tau}-1} \frac{exp\left(-\frac{z}{2\nu_1^2}\right)}{\sqrt{z}}\ \mathrm{d}z\ &= \\ erf\left(\sqrt{\frac{1}{2\nu_1^2}\left(\frac{x}{\tau}-1\right)}\right),\ x \geq \tau, \end{aligned} \tag{6.37}$$

where $erf(\zeta) = \frac{2}{\sqrt{\pi}} \int_0^\zeta e^-t^2\ \mathrm{d}t$ is the error function.

Substituting (6.30) and (6.37) into (6.34) yields

$$\begin{aligned} Pr\{X/Y > \tau\} = \int_{x=\tau}^{\infty} erf\left(\sqrt{\frac{1}{2\nu_1^2}\left(\frac{x}{\tau}-1\right)}\right) \\ \frac{1}{\pi\sqrt{\nu_0^2\nu_2^2-\nu_1^4}} K_0\left(\frac{|x|\nu_0\nu_2}{\nu_0^2\nu_2^2-\nu_1^4}\right) \\ \left\{exp\left(\frac{-x\nu_1^2}{\nu_0^2\nu_2^2-\nu_1^4}\right) + exp\left(\frac{x\nu_1^2}{\nu_0^2\nu_2^2-\nu_1^4}\right)\right\}\ \mathrm{d}x. \end{aligned} \tag{6.38}$$

Making a number of changes of variables in (6.38) yields

$$\begin{aligned} Pr\{(\psi_\sigma n) > \tau\} = Pr\{X/Y > \tau\} = \frac{\sqrt{2}\tau}{\pi} \int_{a=0}^{\infty} erf(\sqrt{a}) \\ .K_0\left(\left|\sqrt{3}a\tau + \frac{2\sqrt{3\pi}\sigma^3\tau}{q^2}\right|\right) \\ .\left\{exp\left(-a\tau - \frac{2\sqrt{\pi}\sigma^3\tau}{q^2}\right) + exp\left(a\tau + \frac{2\sqrt{\pi}\sigma^3\tau}{q^2}\right)\right\}\ \mathrm{d}a. \end{aligned} \tag{6.39}$$

Then, it is easily shown that the probability of false impulse detection in noise using the scale-normalized smoothed singularity index $(\psi_{\sigma,norm}n) = \sigma^2(\psi_\sigma n)$ is

$$Pr\{(\psi_{\sigma,norm}n) > \tau\} = \frac{\sqrt{2}}{\pi}\frac{\tau}{\sigma^2}\int_{a=0}^{\infty} erf(\sqrt{a}) \\ .K_0\left(\left|\frac{\sqrt{3}a\tau}{\sigma^2} + \frac{2\sqrt{3\pi}\tau\sigma}{q^2}\right|\right) \\ .\left\{exp\left(-\frac{a\tau}{\sigma^2} - \frac{2\sqrt{\pi}\tau\sigma}{q^2}\right) + exp\left(\frac{a\tau}{\sigma^2} + \frac{2\sqrt{\pi}\tau\sigma}{q^2}\right)\right\} \, \mathrm{d}a. \tag{6.40}$$

We also compare the false alarm rates of the singularity index with the scale-normalized smoothed second derivative operator

$$(\Theta_{\sigma,norm}f)(t) = \sigma^{1.5}(\Theta_\sigma f)(t), \tag{6.41}$$

where

$$(\Theta_\sigma f)(t) = |(f(t) * g_\sigma(t))''|. \tag{6.42}$$

The operator Θ_σ is the simplest nominal impulse detector, which responds strongly to impulses. It produces a minimal response to the center of an edge, but a large response in the edge neighborhood as discussed in chapter 5. Normalization by $\sigma^{1.5}$ ensures that the maximizing scale matches the impulse width. Note that the scale normalized second derivative index $(\Theta_{\sigma,norm})$ is precisely equivalent to Lindeberg's *scale-normalized maximum eigenvalue of the Hessian* $M_{\gamma-norm}L$ ridge strength measure for a gaussian ridge profile (equation (46) in [60]).

In the noise only scenario, $(\Theta_\sigma n) = |n''_\sigma|$. Since n''_σ is a zero-mean WSS gaussian random process with variance ν_2^2, the PDF of Θ_σ is the half-normal distribution

$$f_{\Theta_\sigma}(\theta) = \sqrt{\frac{2}{\pi}}\frac{1}{\nu_2} exp\left(-\frac{\theta^2}{2\nu_2^2}\right),\ \theta > 0. \tag{6.43}$$

Hence,

$$\begin{aligned} Pr\{(\Theta_\sigma n) > \tau\} &= 1 - Pr\{(\Theta_\sigma n) \leq \tau\} \\ &= 1 - \int_0^\tau \sqrt{\frac{2}{\pi}}\frac{1}{\nu_2} exp\left(-\frac{\theta^2}{2\nu_2^2}\right)\ \mathrm{d}\theta \\ &= 1 - erf\left(\frac{\tau}{\sqrt{2}\nu_2}\right), \end{aligned} \tag{6.44}$$

which is the probability of false impulse detection when the operator Θ_σ is applied to a signal comprising of only a zero-mean WSS gaussian random process. Then, the probability of false impulse detection in the noise only scenario by the scale-normalized smoothed second derivative operator is

$$Pr\{(\Theta_{\sigma,norm} n) > \tau\} = 1 - erf\left(\frac{\tau}{\sqrt{2}\nu_2\sigma^{1.5}}\right). \tag{6.45}$$

Note that the probabilities in (6.40) and (6.45) have well-behaved integrals that can be easily evaluated numerically using packages such as the Matlab library.

6.3.2 True Impulse Detection in the Presence of Noise

We now study the detection power of the new index by finding the probability of detecting a noise corrupted impulse signal (situated, without loss of generality at the origin) when the scale-normalized smoothed singularity index (6.3) (with $\gamma = 1$) is applied. Let $f(t)$ be the noise corrupted impulse signal: $f(t) = s(t) + n(t)$, where $s(t)$ is the deterministic signal and $n(t)$ is a zero-mean WSS gaussian random noise process as before. We assume $s(t)$ to be a gaussian smoothed impulse of height K and scale w as before, i.e. $s(t) = K\delta_w(t)$. Let $s_\sigma(t) = s(t) * g_\sigma(t)$, $s'_\sigma(t) = (s(t) * g_\sigma(t))'$, and $s''_\sigma(t) = (s(t) * g_\sigma(t))''$. Then the response of the scale normalized smoothed singularity index (6.3) applied to the noise corrupted impulse signal $f(t) = s(t) + n(t)$ may be expressed

$$\begin{aligned}(\psi_{\sigma,norm} f)(t) &= \sigma^2 \frac{|(s_\sigma(t) + n_\sigma(t))(s_\sigma(t) + n_\sigma(t))''|}{1 + [(s_\sigma(t) + n_\sigma(t))']^2} \\ &= \frac{|\sigma^2 (s_\sigma(t) + n_\sigma(t))(s''_\sigma(t) + n''_\sigma(t))|}{1 + [(s'_\sigma(t) + n'_\sigma(t))]^2}.\end{aligned} \tag{6.46}$$

Since $s(t)$ is a deterministic signal and $n_\sigma(t)$, $n'_\sigma(t)$, and $n''_\sigma(t)$ are all zero-mean gaussian random processes, we have: $E[s_\sigma(t) + n_\sigma(t)] = s_\sigma(t)$, $var[s_\sigma(t) + n_\sigma(t)] = var[n_\sigma(t)] = \nu_0^2$, $E[s'_\sigma(t) + n'_\sigma(t)] = s'_\sigma(t)$, $var[s'_\sigma(t) + n'_\sigma(t)] = var[n'_\sigma(t)] = \nu_1^2$, $E[s''_\sigma(t) + n''_\sigma(t)] = s''_\sigma(t)$, and $var[s''_\sigma(t) + n''_\sigma(t)] = var[n''_\sigma(t)] = \nu_2^2$. Further, $s_\sigma(t)$, $s'_\sigma(t)$, and $s''_\sigma(t)$ are given by:

$$\begin{aligned}s_\sigma(t) &= \frac{Kw}{\sqrt{\sigma^2 + w^2}} e^{\frac{-t^2}{2(\sigma^2 + w^2)}} \\ s'_\sigma(t) &= \frac{-Kwt}{(\sigma^2 + w^2)^{3/2}} e^{\frac{-t^2}{2(\sigma^2 + w^2)}} \\ s''_\sigma(t) &= \frac{Kw}{(\sigma^2 + w^2)^{3/2}} \left(\left(\frac{t^2}{\sigma^2 + w^2} - 1 \right) e^{\frac{-t^2}{2(\sigma^2 + w^2)}} \right).\end{aligned} \tag{6.47}$$

At $t = 0$,

$$s_\sigma(0) = \frac{Kw}{\sqrt{\sigma^2 + w^2}}, \; s'_\sigma(0) = 0, \; s''_\sigma(0) = \frac{-Kw}{(\sigma^2 + w^2)^{3/2}}. \tag{6.48}$$

Denote the random variable $(\psi_{\sigma,norm} f)(0)$ by the ratio $X_{norm}/Y = \sigma^2 X/Y$, where again X and Y are statistically independent. The random variables $(s'_\sigma(0) + n'_\sigma(0))$ is zero-mean gaussian with variance ν_1^2, since $s'_\sigma(0) = 0$. Hence, the PDF of the variable $Y = 1 + [(s'_\sigma(0) + n'_\sigma(0))]^2$ is as defined in (6.27).

Further, $(s_\sigma(0) + n_\sigma(0))$ and $(s''_\sigma(0) + n''_\sigma(0))$ are non-zero-mean, correlated gaussian random variables having means $s_\sigma(0)$ and $s''_\sigma(0)$, variances ν_0^2 and ν_2^2, respectively, and correlation coefficient $\rho = \frac{-\nu_1^2}{\nu_0 \nu_2}$. Unlike the zero-mean case, there exists no simple closed form expression for the PDF of the product $(s_\sigma(0) + n_\sigma(0))(s''_\sigma(0) + n''_\sigma(0))$.

Instead, let $X_1 = |(s_\sigma(0) + n_\sigma(0))|$ and $X_2 = |(s''_\sigma(0) + n''_\sigma(0))|$. Hence, the random variable $X = X_1 X_2$ is the product of two non-zero mean folded normal random processes with means $s_\sigma(0)$ and $s''_\sigma(0)$, variances ν_0^2 and ν_2^2, respectively, and correlation coefficient ρ. The PDF of the product $X = X_1 X_2$ can be determined as follows [80]:

$$f_X(x) = \int_0^\infty \frac{1}{|u|} f_{X_1 X_2}\left(u, \frac{x}{u}\right) \, \mathrm{d}u, \; x > 0, \tag{6.49}$$

where the joint PDF $f_{X_1 X_2}$ is the bivariate folded normal density function defined in (3.1) in [83]. Hence, the PDF of the numerator X_{norm} is

$$f_{X_{norm}}(x) = \frac{1}{\sigma^2} f_X\left(\frac{x}{\sigma^2}\right), x > 0. \tag{6.50}$$

Given a threshold $\tau > 0$, the probability of detecting an impulse at the origin is

$$\begin{aligned} Pr\{(\psi_{\sigma,norm} f)(0) > \tau\} = \int_{x=\tau}^\infty erf\left(\sqrt{\frac{1}{2\nu_1^2}\left(\frac{x}{\tau} - 1\right)}\right) \\ f_{X_{norm}}(x) \, \mathrm{d}x, \end{aligned} \tag{6.51}$$

which is arrived at by following the sequence of steps from (6.32) to (6.37).

Finally, for comparison, the probability of detecting an impulse at the origin using the scale-normalized smoothed second derivative operator $(\Theta_{\sigma,norm} f)(0) = \sigma^{1.5}(\Theta_\sigma f)(0)$ is

$$Pr\{(\Theta_{\sigma,norm} f)(0) > \tau\} = 1 - \int_0^\tau f_{\Theta_{\sigma,norm}}(\theta)\ \mathrm{d}\theta, \tag{6.52}$$

where

$$\begin{aligned} f_{\Theta_{\sigma,norm}}(\theta) &= \frac{1}{\sqrt{2\pi}\nu_2\sigma^{1.5}} \\ \cdot \left[e^{-\frac{\left(\frac{\theta}{\sigma^{1.5}} - s''_\sigma(0)\right)^2}{2\nu_2^2}} + e^{-\frac{\left(\frac{\theta}{\sigma^{1.5}} + s''_\sigma(0)\right)^2}{2\nu_2^2}} \right] &, \theta > 0. \end{aligned} \tag{6.53}$$

The probabilities given by (6.51) and (6.52) can again be easily evaluated numerically.

6.3.3 Simulations

The detection power and the false alarm rates of the scale-normalized singularity index and the second derivative operator were assessed using receiver operating characteristic (ROC) curves. Our simulations assumed gaussian smoothed impulses $s(t) = K\delta_w(t)$, of three different heights K (5, 10, and 20), each submerged in white gaussian noise of four different variances q^2 (5, 10, 50, and 100). The scale w of each impulse was fixed at 1.5. The scale σ of the gaussian low pass filter was set to match the impulse scale i.e. $\sigma = w = 1.5$, since we are using the scale-normalized versions of the index

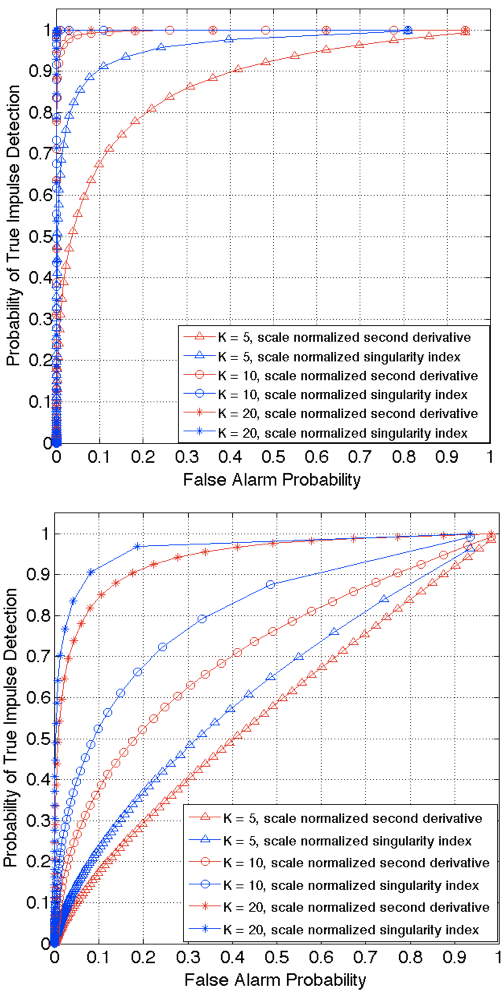

Figure 6.2: ROC curves plotting the probability of true impulse detection for gaussian smoothed impulses of different heights against the false alarm probability for the scale-normalized singularity index (blue) and the second derivative operator (red). Top row: $q^2 = 5$ (left) and $q^2 = 10$ (right). Bottom row: $q^2 = 50$ (left) and $q^2 = 100$ (right). The scale w of the impulses and the scale σ of the smoothing gaussian were both fixed at 1.5.

and the second derivative operator. We next describe the threshold selection mechanism that was used to generate the ROC curves.

For the scale-normalized smoothed singularity index, the scale-space peak response (at the origin) to a gaussian smoothed impulse in the absence of noise is scale invariant and is given by $(\psi_{\sigma=w,norm}s)(0) = \frac{K^2}{4}$ (see chapter 5). The threshold τ was varied in equal increments (set to $0.05 \times \frac{K^2}{4}$) from 0.05 to $M \times \frac{K^2}{4}$, where M is an integer that was set to 6 to yield a sufficiently large range of thresholds. Note that the range of thresholds is different for different values of K. The same thresholds were used for computing both the false alarm and true impulse detection probabilities of the scale-normalized smoothed singularity index using (6.40) and (6.51), respectively. The false alarm and true impulse detection probabilities were then subsequently used to generate the ROC curves.

Similarly, for the scale-normalized smoothed second derivative operator, the scale-space peak response to a gaussian smoothed impulse in the absence of noise is given by $(\Theta_{\sigma=w,norm}s)(0) = \frac{K}{2^{3/2}\sqrt{w}}$ (see chapter 5), which is not scale invariant. Here again, the threshold τ was varied in equal increments (set to $0.05 \times \frac{K}{2^{3/2}\sqrt{w}}$) from 0.05 to $M \times \frac{K}{2^{3/2}\sqrt{w}}$, where $M = 6$. The false alarm and true impulse detection probabilities of the scale-normalized second derivative operator were computed for the same thresholds using (6.45) and (6.52), respectively, which were then used to generate the ROC curves.

Fig. 6.2 plots the ROC curves for the scale-normalized smoothed singularity index and the second derivative operator that were obtained using the threshold selection mechanism described above for gaussian smoothed impulses of heights $K = 5$, 10, and 20, scale $w = 1.5$, and each submerged in noise of variances $q^2 = 5$, 10, 50, and 100. The superior performance of the singularity index over the second derivative operator is immediately apparent from Fig. 6.2. At low noise variances (e.g., as depicted in the plot for $q^2 = 5$ in Fig. 6.2), the difference in ROC curves is larger for smaller values of K (e.g., $K = 5$). This accords with intuition, since one would expect that as K increases in the presence of noise having low variance, the detection power of the singularity index and the second-derivative operator would be similar. On the other hand, as the noise variance increases, the difference in ROC curves between the scale-normalized smoothed singularity index and the second derivative operator increases with K (e.g., as depicted in the plot for $q^2 = 100$ in Fig. 6.2). The ability of both the singularity index and the second derivative operator to reliably detect small impulses (e.g., $K = 5$) while at the same time keeping the false alarm rates low diminishes in the presence of noise having large variance. Yet, the singularity index outperforms the second derivative operator even for small impulses submerged in noise of large variance.

We also considered gaussian smoothed impulses whose scales w were different, but had the same height. Three values of w (1.5, 3, and 4.5) were considered, while K and q^2 were fixed at 10 and 100, respectively. The scale σ of the gaussian low pass filter was set to match the impulse scale w. Fig. 6.3 plots the ROC curves for the scale-normalized smoothed singularity index and

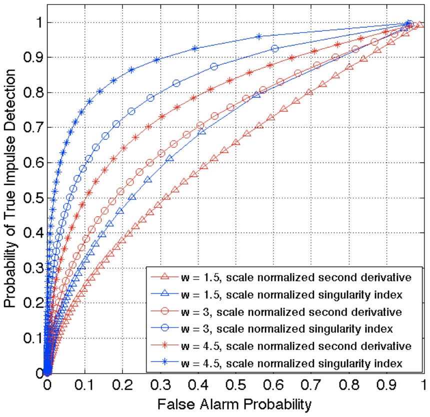

Figure 6.3: ROC curves plotting the probability of true impulse detection for gaussian smoothed impulses of different scales w against the false alarm probability for the scale-normalized singularity index (blue) and the second derivative operator (red). The plots were generated for $K = 10$, $q^2 = 100$ and $\sigma = w$.

the second derivative operator using the same threshold selection mechanism as before for the gaussian smoothed impulses of different scales. Again, the singularity index offers superior performance when compared to the second-derivative operator. What is revealing is that both the operators offer better detection power at lower false alarm rates when the scale w of the impulse increases. This is due to the fact that since the scale σ of the gaussian low pass filter is set to match the impulse scale w, the increase in σ reduces the effect of noise and yields better ROC curves. Note that as described earlier, in the absence of noise, the scale-space peak response to a gaussian smoothed impulse of scale-normalized singularity index is scale invariant. However, this does not hold when the impulse is submerged in noise.

6.4 Edge Suppression in the Presence of Noise

The smoothed singularity index defined in (6.2) is designed to be edge suppressing. Hence, we consider the problem of finding the probability of incorrectly detecting a noise corrupted edge at the origin when the scale-normalized smoothed singularity index defined in (6.3) ($\gamma = 1$) is applied. As in the impulse detection scenario, let $f(t) = s(t) + n(t)$. We model the edge signal $s(t)$ as a Heaviside step $u(t)$ of height $K > 0$ smoothed by a gaussian $g_w(t)$: $s(t) = K\Phi\left(\frac{t}{w}\right) - \frac{K}{2}$, where $\Phi\left(\frac{t}{w}\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{t}{w}} e^{\frac{-x^2}{2}}\, dx$. The constant $\frac{K}{2}$ is subtracted to ensure that the edge height is 0 at the origin and symmetric about it. Then, $s_\sigma(t)$, $s'_\sigma(t)$, and $s''_\sigma(t)$ are given by:

$$\begin{aligned} s_\sigma(t) &= K\Phi\left(\frac{t}{\sqrt{\sigma^2 + w^2}}\right) - \frac{K}{2} \\ s'_\sigma(t) &= \frac{K}{\sqrt{2\pi(\sigma^2 + w^2)}} e^{-\frac{-t^2}{2(\sigma^2+w^2)}} \\ s''_\sigma(t) &= \frac{-Kt}{\sqrt{2\pi}(\sigma^2 + w^2)^{3/2}} e^{-\frac{-t^2}{2(\sigma^2+w^2)}}. \end{aligned} \tag{6.54}$$

At $t = 0$,

$$s_\sigma(0) = 0, \; s'_\sigma(0) = \frac{K}{\sqrt{2\pi(\sigma^2 + w^2)}}, \; s''_\sigma(t) = 0. \tag{6.55}$$

Again, denote the random variable $(\psi_{\sigma,norm} f)(0)$ as the ratio $X_{norm}/Y = \sigma^2 X/Y$, where X and Y are statistically independent. The PDF of X_{norm} is then

$$\begin{aligned} f_{X_{norm}}(x) = \frac{1}{\pi\sqrt{\nu_0^2\nu_2^2 - \nu_1^4}\sigma^2} K_0\left(\frac{|\frac{x}{\sigma^2}|\nu_0\nu_2}{\nu_0^2\nu_2^2 - \nu_1^4}\right) \\ \cdot\left\{exp\left(\frac{\frac{-x}{\sigma^2}\nu_1^2}{\nu_0^2\nu_2^2 - \nu_1^4}\right) + exp\left(\frac{\frac{x}{\sigma^2}\nu_1^2}{\nu_0^2\nu_2^2 - \nu_1^4}\right)\right\}, \; x > 0, \end{aligned} \tag{6.56}$$

since X_{norm} is the absolute value of the product of two correlated, zero-mean gaussian random variables. The PDF of the random variable Y is

$$f_Y(y) = P(y-1),\ y > 1, \tag{6.57}$$

where

$$P(y) = \frac{1}{2\sqrt{2\pi(y)}\nu_1}\left[e^{-\frac{\left(\sqrt{y}-s'_\sigma(0)\right)^2}{2\nu_1^2}} + e^{-\frac{\left(\sqrt{y}+s'_\sigma(0)\right)^2}{2\nu_1^2}}\right]. \tag{6.58}$$

Then, given a threshold $\tau > 0$, the probability of detecting an edge at the origin is

$$\begin{aligned} Pr\{(\psi_{\sigma,norm}f)(0) > \tau\} = \int_{x=\tau}^{\infty}\left(\int_{z=0}^{\frac{x}{\tau}-1} P(z)\ \mathrm{d}z\right) \\ f_{X_{norm}}(x)\ \mathrm{d}x. \end{aligned} \tag{6.59}$$

Finally, the probability of detecting an edge at the origin by the scale-normalized second derivative operator is

$$Pr\{(\Theta_{\sigma,norm}f)(0) > \tau\} = 1 - erf\left(\frac{\tau}{\sqrt{2}\nu_2\sigma^{1.5}}\right), \tag{6.60}$$

which is independent of the edge height K and scale w. This follows since the random process $(s''_\sigma(0) + n''_\sigma(0))$ is zero-mean gaussian with variance ν_2^2. In fact, for the second derivative operator, the probability of detecting an edge is identical to the probability of finding an impulse (6.45) in the noise-only scenario!

6.4.1 Simulations

Since both the singularity index and the second derivative operator are, by design, edge suppressing, we treat the detection of an edge in the presence of noise as being identical to the false alarm scenario, i.e, detection of an impulse singularity when none is present. Hence, we plot ROC curves using the same threshold selection mechanism described in the previous section. Here again, we consider gaussian smoothed impulses of three different heights K (5, 10, and 20), each submerged in white gaussian noise of four different variances q^2 (5, 10, 50, and 100). The scale w of each impulse was fixed at 1.5 and the scale σ of the gaussian low pass filter was set to match the impulse scale.

Fig. 6.4 plots the ROC curves for the scale-normalized smoothed singularity index and the second derivative operator for gaussian smoothed impulses of heights $K = 5$, 10, and 20, scale $w = 1.5$, and each submerged in noise of variances $q^2 = 5$, 10, 50, and 100. Notice in Fig. 6.4 that as expected, the ROC curves for the scale-normalized second derivative operator are identical to the ROC curves plotted in Fig. 6.2. It is evident from Fig. 6.4 that the singularity index clearly outperforms the second derivative operator across the four different noise variances in detecting true impulses while maintaining low false alarm rates arising out of edge detection.

6.5 2-D Singularity Index

The singularity index is readily extended to 2-D. As seen in chapter 5, the 2-D counterparts for (6.1) and (6.2) are:

$$(\psi f)(x, y) = \frac{|f(x, y)||\nabla^2 f(x, y)|}{1 + |\nabla f(x, y)|^2}, \tag{6.61}$$

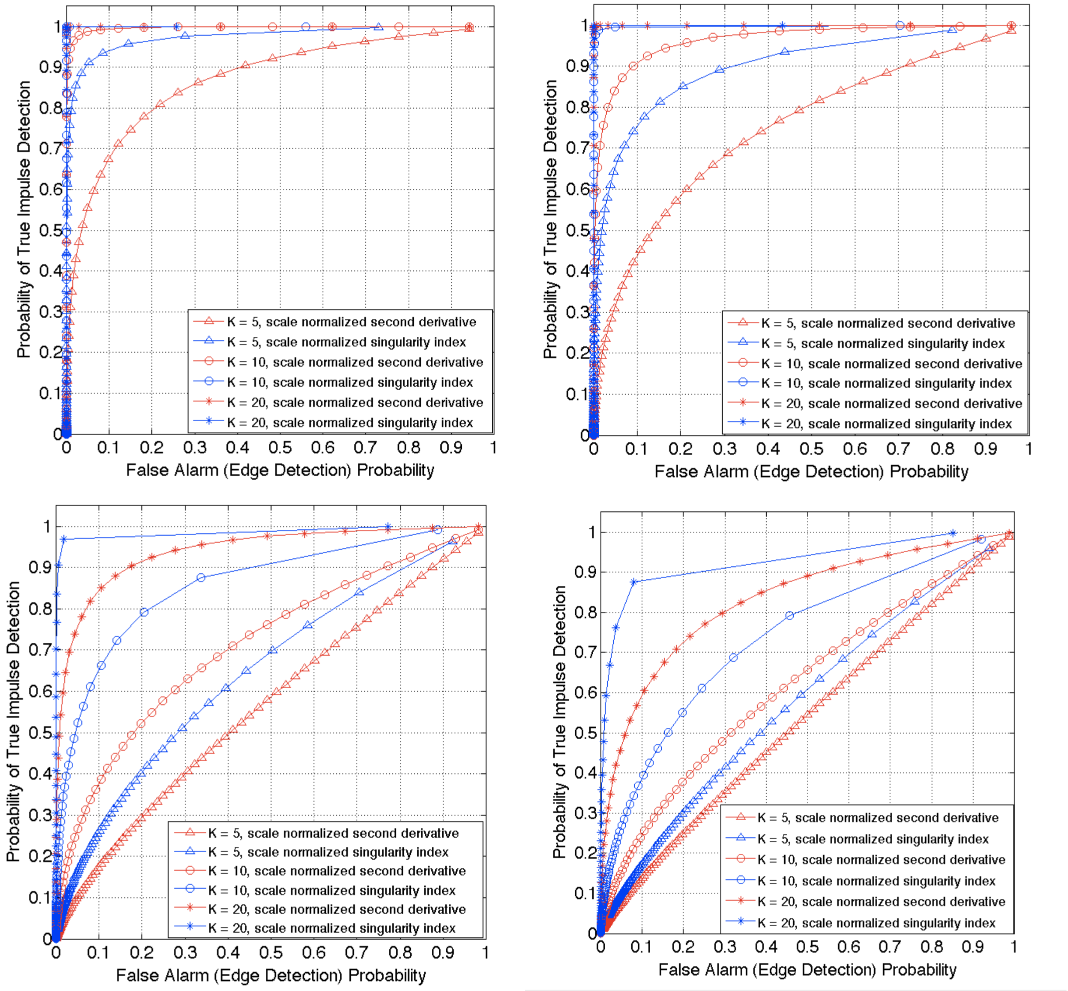

Figure 6.4: ROC curves plotting the probability of true impulse detection for gaussian smoothed impulses of different heights against the false alarm probability arising out of edge detection for the scale-normalized singularity index (blue) and the second derivative operator (red). Top row: $q^2 = 5$ (left) and $q^2 = 10$ (right). Bottom row: $q^2 = 50$ (left) and $q^2 = 100$ (right). The scale w of the impulses and the scale σ of the smoothing gaussian were both fixed at 1.5.

and

$$(\psi_\sigma f)(x,y) = \frac{|g_\sigma(x,y) * f(x,y)||\nabla^2 g_\sigma(x,y) * f(x,y)|}{1 + |\nabla g_\sigma(x,y) * f(x,y)|^2}, \tag{6.62}$$

where $f(x,y), R^2 \rightarrow R$ is a 2-D signal or function, the 2-D Laplacian operator $\nabla^2 f(x,y) = \left(\frac{\partial^2 f(x,y)}{\partial x^2} + \frac{\partial^2 f(x,y)}{\partial y^2}\right)$, the 2-D gradient operator $\nabla f(x,y) = \left(\frac{\partial f(x,y)}{\partial x}, \frac{\partial f(x,y)}{\partial y}\right)^T$, and $g(x,y)$ is a 2-D unit area isotropic smoothing gaussian $g(x,y) = \frac{1}{2\pi\sigma^2} e^{\frac{-(x^2+y^2)}{2\sigma^2}}$. In particular, if $f(x,y)$ represents the image brightness function, then the 2-D singularity index (6.62) is useful for detecting impulse masses in images.

It is often of interest to detect curvilinear masses in images, which have a dominant orientation along which the second order directional derivative attains a local extremum. Hence, as described in chapter 5, the index (6.62) is modified to account for the dominant orientation

$$(\psi_\sigma f)(x,y) = \frac{|f_{0,\theta_D,\sigma}(x,y) f_{2,\theta_D,\sigma}(x,y)|}{1 + |f_{1,\theta_D,a\sigma}(x,y)|^2}, \tag{6.63}$$

where $f_{0,\theta_D,\sigma}(x,y)$, $f_{1,\theta_D,a\sigma}(x,y)$, and $f_{2,\theta_D,\sigma}(x,y)$ are the responses to the zero, first and second order isotropic gaussian derivative filters along the direction specified by $\theta_D(x,y)$ and at scale σ. Estimation of $\theta_D(x,y)$ (the direction along which the second order derivative attains a local extremum) is straightforward and is achieved by exploiting the steerable property of isotropic gaussian directional derivatives as described in [29]. Further, the first derivative response in (6.63) is computed at a scale $a\sigma$ to attenuate edge side-lobe responses as discussed in chapter 5.

The scale-normalized 2-D singularity index is

$$\begin{aligned}(\psi_{\sigma,norm}f)(x,y) &= \sigma^2(\psi_\sigma f)(x,y)\\ &= \sigma^2\frac{|f_{0,\theta_D,\sigma}(x,y)f_{2,\theta_D,\sigma}(x,y)|}{1+|f_{1,\theta_D,a\sigma}(x,y)|^2},\end{aligned} \tag{6.64}$$

where σ is the scale of the isotropic smoothing gaussian. The scale-normalized 2-D index (6.64) can detect curvilinear masses at multiple scales. This is achieved by varying the scale σ of the isotropic smoothing gaussian and computing the maximum scale-normalized index response across all scales. Once the maximum response is computed at every location (x, y), non-maxima suppression (NMS) is applied along the dominant orientation $\theta_D(x, y)$. This design is similar to Canny's for directional edge detection [20].

We used the results of the 1-D analyses to guide the choice of threshold values for the 2-D scale-normalized singularity index $(\psi_{\sigma,norm}f)(x, y)$ and the second derivative operator responses $\sigma^{1.5}|f_{2,\theta_D,\sigma}(x, y)|$ computed on noisy images. The rationale behind using the 1-D analyses directly in the 2-D scenario is as follows: model a smoothed 2-D line impulse profile by a 2-D isotropic gaussian of strength $K > 0$ and scale σ convolved with the impulse sheet $\delta(x)$:

$$\begin{aligned}f(x,y) &= \frac{K}{\sqrt{2\pi}\sigma}e^{\frac{-(x^2+y^2)}{2\sigma^2}} * \delta(x)\\ &= \frac{K}{\sqrt{2\pi}\sigma}e^{\frac{-x^2}{2\sigma^2}}\int_{-\infty}^{\infty} e^{\frac{-v^2}{2\sigma^2}}\,\mathrm{d}v = Ke^{\frac{-x^2}{2\sigma^2}}, \forall y\end{aligned} \tag{6.65}$$

Then, the conclusions regarding the response of the 2-D singularity index (6.61) to the 2-D line impulse profile are the same as that of the 1-D singularity index (6.1) for the 1-D impulse profile. Note that the analysis of the detection power and the false alarm rates of the 2-D singularity index applied to a 2-D white gaussian noise process is possible, though tedious.

To determine the threshold values, we used the 1-D gaussian smoothed impulse model of height $K = 20$ and scale $w = 3$ submerged in a zero-mean WSS gaussian random noise process of variance $q^2 = 100$. The scale σ of the smoothing gaussian was set equal to w. We generated ROC curves for the 1-D scale normalized singularity index and the second derivative operators by plotting the probabilities of true impulse detection against the false alarm probabilities using the threshold selection mechanism described before. The two ROC curves are illustrated in Fig. 6.5. Clearly, the performance of the scale-normalized singularity index is superior to the performance of the second derivative operator on evidence of the ROC curves. On each ROC curve, we considered an operating point (denoted by the triangles in Fig. 6.5) where the probability of true impulse detection was approximately 0.65. The threshold values for the singularity index and the second derivative operator were based on these operating points and were set to 55.05 and 3.52, respectively.

Figs. 6.6 and 6.7 show examples of pristine and noisy images (rows 1 and 2, respectively). The noisy images were generated by contaminating the pristine images with independent, zero mean additive white gaussian noise (AWGN) of variance 100. The 2-D scale-normalized singularity index (6.63) and the scale-normalized second-derivative operator $\sigma^{1.5}|f_{2,\theta,\sigma}(x, y)|$ responses (with NMS) were computed over 5 scales on the noisy images illustrated in Figs. 6.6 and 6.7. It is worth reiterating here that the scale-normalized

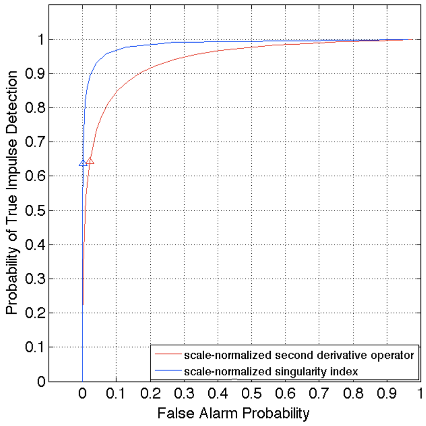

Figure 6.5: Receiver operating characteristic curves for the scale-normalized 1-D singularity index (blue) and the second derivative operator (red) computed for an impulse of height $K = 20$ and scale $w = 3$ submerged in noise of variance $q^2 = 100$. The scale σ of the smoothing gaussian was also set to 3. The triangles denote the operating points on each curve corresponding to the chosen threshold values.

second-derivative operator $\sigma^{1.5}|f_{2,\theta,\sigma}(x,y)|$ corresponds exactly to Lindeberg's $M_{\gamma-norm}L$ ridge strength measure defined in (46) in [60]. The lowest scale σ of the isotropic gaussian was set to 3 pixels, while each subsequent coarser scale was larger than the previous finer scale by a factor of $\sqrt{2}$. The NMS responses were then thresholded to detect curvilinear masses using threshold values 55.05 and 3.52 for the singularity index and the second derivative operator, respectively.

The results of the thresholding are shown in rows 3 and 4 for the singularity index and the second derivative operator, respectively. It is evident from row 3 in Figs. 6.6 and 6.7 that the singularity index detects most of the salient curvilinear structures with good contour continuity while at the same time suppressing edges. On the other hand, the second derivative operator results in a lot more clutter and in the detection of edges as well. The detection of edges by the second derivative operator should not be surprising, since as shown in our analyses, the probability of falsely detecting an impulse in the noise only scenario is exactly the same as the probability of incorrectly detecting an edge suppressed in noise.

6.6 Summary

This chapter has developed detailed theoretical analyses of the detection power and false alarm probabilities of the new 1-D singularity index that was recently designed for impulse detection in signals of arbitrary dimensionality and presented in chapter 5. By design, the singularity index amplifies response to impulses, while at the same time delivering powerful attenuation to edges. Our theoretical analyses and subsequent simulations involving ROC curves and experiments with real images corroborate this claim. By comparison, the simple second derivative operator, which is a nominal impulse detector, fails in reliably detecting impulses and suppressing edges in a noisy environment. The index is naturally scalable and is computationally efficient, since it exploits the steerable property of isotropic gaussian derivatives. We next show how the singularity index is employed in a novel stereo model for disparity estimation on stereo mammographic images.

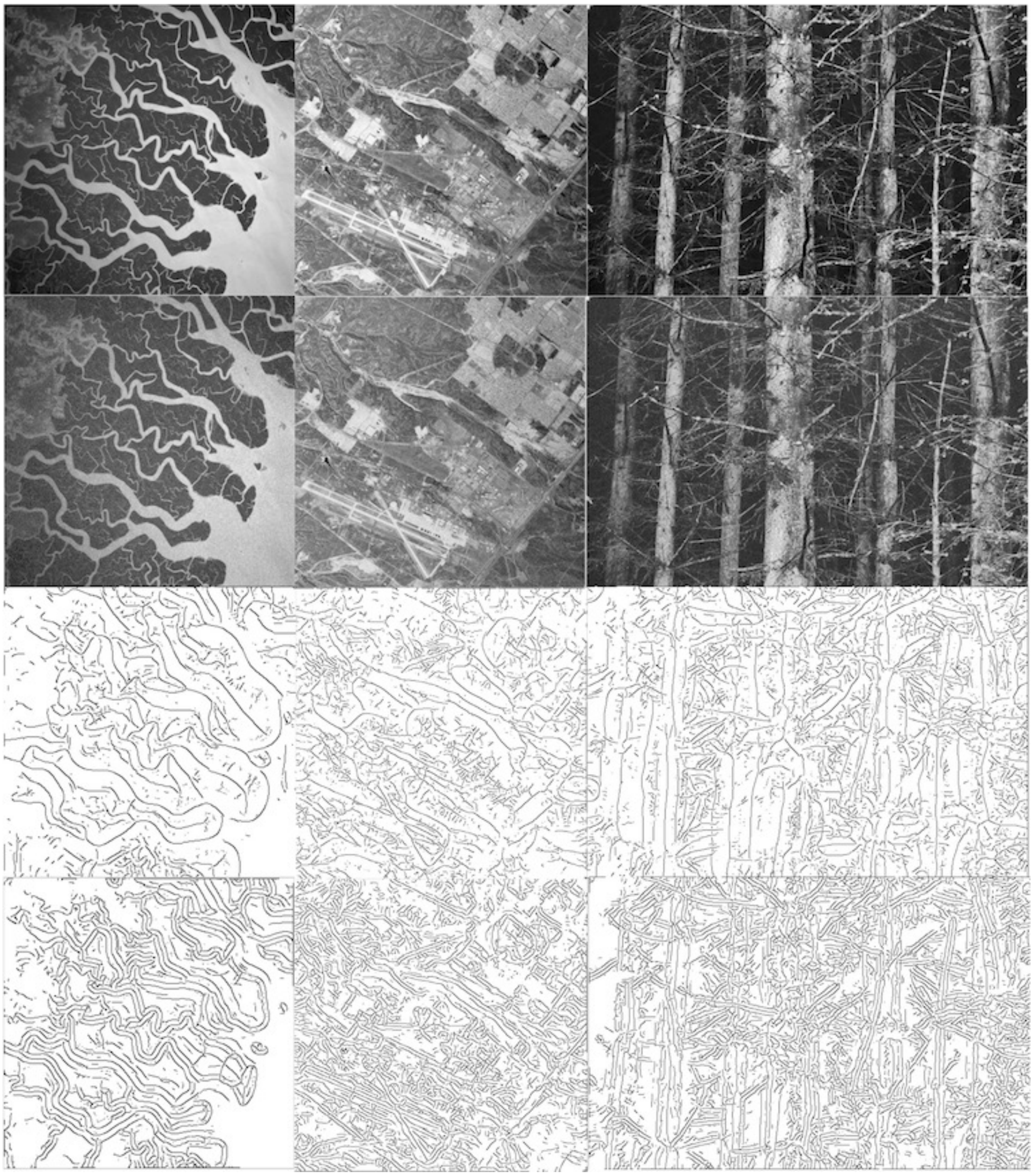

Figure 6.6: Row 1: Original images (column 1: The Ganges river delta, NASA, courtesy of nasaimages.org, column 2: An aerial image (courtesy University of Southern California), column 3: An image of pine tree trunks), row 2: Corresponding noisy images (AWGN of variance 100), rows 3 and 4: Curvilinear structures detected on thresholding the NMS scale-normalized singularity index response and the NMS scale-normalized second derivative operator response, respectively.

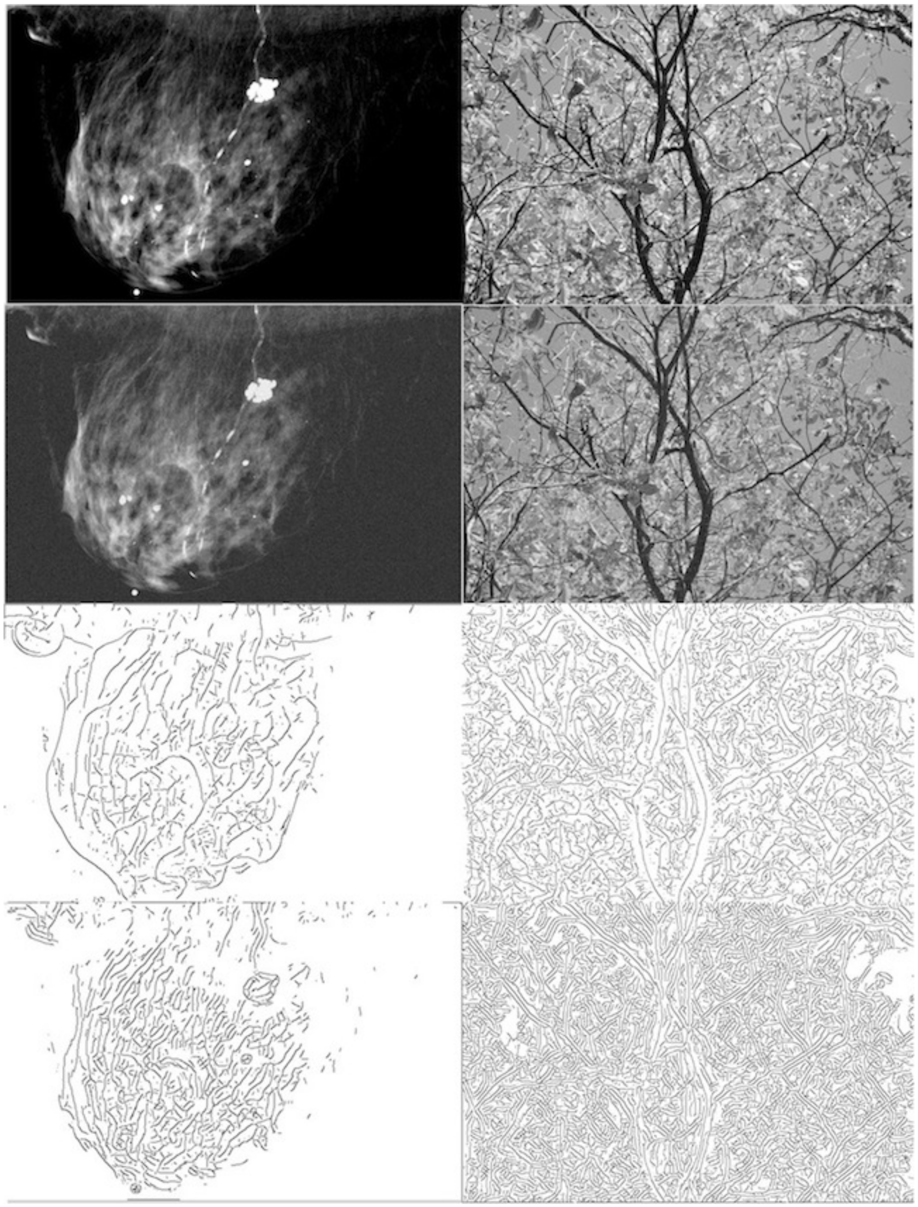

Figure 6.7: Row 1: Original images (column 1: A mammogram courtesy Emory University, Atlanta, GA, column 2: A foliage scene), row 2: Corresponding noisy images (AWGN of variance 100), rows 3 and 4: Curvilinear structures detected on thresholding the NMS scale-normalized singularity index response and the NMS scale-normalized second derivative operator response, respectively.

Chapter 7

Disparity Estimation on Stereo Mammograms

7.1 Introduction

In this chapter, we describe a novel computational stereo model that is specifically directed towards estimating a dense disparity map from a pair of stereo mammograms [70]. Our proposed model employs the new singularity index described in chapter 5 to better estimate disparity along the singular points arising out of a dense network of curvilinear structures evident on stereo mammograms. We begin this chapter with a description of the canonical stereo model in section 7.2, which serves as the baseline for the proposed model. We then discus the proposed stereo model in section 7.3, which is followed by a discussion of the optimization strategy in section 7.4. The experimental methodology and results are described in sections 7.5 and 7.6, respectively. We summarize the chapter in section 7.7.

7.2 Canonical Stereo Model

We first describe the canonical stereo model, which we will use as a baseline for comparing the performance of the proposed stereo model. Let $I_1(\mathbf{x})$ and $I_2(\mathbf{x})$ denote the two images of a rectified stereo pair (such that the corresponding rows are registered), where $\mathbf{x} = (x, y)^T$ represents an image pixel location belonging to the discrete rectangular domain Z^2. Let $I_1(\mathbf{x})$ denote the reference image. Let $D(\mathbf{x})$ denote the integer-valued disparity at the image location $\mathbf{x}$. Then, according to the canonical stereo model, D is computed as the minimizer of the following energy functional:

$$E(D(\mathbf{x})) = E_P(D(\mathbf{x})) + \lambda_1 E_S(D(\mathbf{x})), \tag{7.1}$$

where $E_P(D(\mathbf{x}))$ represents the photometric term, $E_S(D(\mathbf{x}))$ represents the disparity smoothness term, and λ_1 is a weight indicating the relative importance of the smoothness term.

The photometric term penalizes disparity assignments that result in large intensity differences between the corresponding points: $E_P(D(\mathbf{x})) = (I_1(x, y) - I_2(x + D(x, y), y))^2$. By penalizing large intensity differences between the corresponding points of a stereo mammogram pair, we implicitly assume that the projected intensities along two linear x-ray trajectories separated by a small angle are similar. The disparity smoothness term penalizes large changes in disparity gradient and is defined as the sum of absolute difference in disparity between that assigned to a given pixel and its immediate neighbors: $E_S(D(\mathbf{x})) = \sum_{\mathbf{p} \in N} |D(\mathbf{x}) - D(\mathbf{p})|$, where $\mathbf{p}$ is a pixel location in the neighborhood N of $\mathbf{x}$.

7.3 Proposed Stereo Model

A major drawback of the canonical stereo model is that the disparity smoothness constraint is enforced everywhere irrespective of the presence of isolated curvilinear masses, object boundaries and occluding surfaces, which invariably result in depth discontinuities. This problem can be addressed by regulating the smoothness term such that disparity smoothness is de-emphasized at certain image locations, e.g. at edges between objects and surface boundaries (e.g., as in [65] and [66]). In a stereo mammogram, depth discontinuities arise from a complex pattern of occluding fine scale curvilinear structures that need to be preserved in the disparity space. Our proposed stereo model is geared towards addressing this issue.

Depth discontinuities are often co-located with the intensity discontinuities that are observed in input stereo image pairs [63]. This property can be exploited to de-emphasize disparity smoothness and to promote curvilinear masses in disparity space to better preserve fine-scale curvilinear structures. Our stereo model is based on this philosophy. The problem of preserving depth discontinuities necessitates the reliable detection of singularity locations in the reference view of the stereo mammogram pair.

Two kinds of singularities are encountered in stereo mammograms: a) impulse singularities, such as those arising from isolated curvilinear masses or the medial axis of fine scale curvilinear structures, and b) edge singularities, such as those arising from the boundaries of curvilinear structures and other anatomical components. Fig. 7.1 shows the reference view of two stereo mammogram pairs (top) with different scan-lines highlighted in color. Also shown in Fig. 7.1 (bottom) are the corresponding 1-D cross-sectional luminance profiles along the different scan-lines. The presence of impulse and edge singularities is clearly evident in the 1-D luminance cross-sectional profiles (Fig. 7.1 (bottom)). Similarly, as shown in Fig. 7.2, impulse and edge singularities are also present in the 1-D cross sectional profiles of disparity as evident in the example disparity map computed using the proposed model for the stereo pair shown in Fig. 7.3. Towards finding these kinds of singularities reliably and robustly in stereo mammogram images, we employ the new singularity index described in chapter 5. For example, Fig. 7.4 shows the response of the 1-D smoothed singularity index

$$\psi_\sigma[f(x)] = \frac{|g_\sigma * f(x)||g''_\sigma * f(x)|}{1 + |g'_{a\sigma} * f(x)|^2} \tag{7.2}$$

when applied to a 1-D luminance cross-sectional profile from a stereo mammogram image. It is evident from Fig. 7.4 that the singularity index responds strongly to isolated impulses and is not sensitive to the polarity of the impulse.

7.3.1 Singularity Index for Disparity Estimation

Recall that the k^{th}-order index defined in chapter 5

$$\psi^k_\sigma[f(x)] = \frac{|g^{k-1}_\sigma * f(x)||g^{k+1}_\sigma * f(x)|}{1 + |g^k_{a\sigma} * f(x)|^2}, \tag{7.3}$$

is a generalization of (7.2), where ($k = 1$) yields (7.2), an impulse index, and $k = 2$ yields an edge singularity index that responds strongly to edges, which may arise from important breast structures such as large masses and

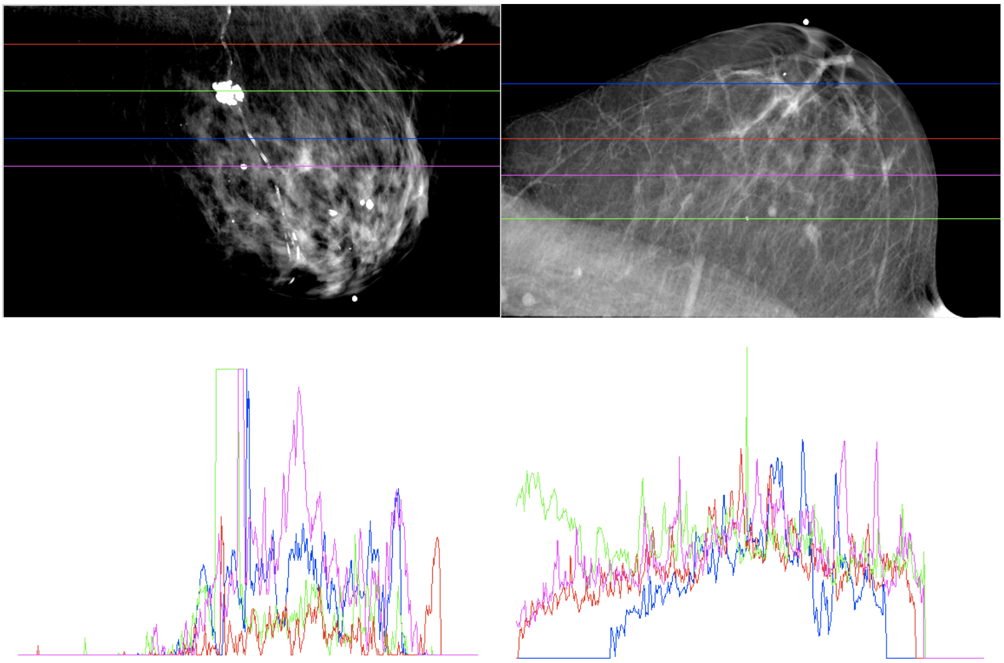

Figure 7.1: Reference views of two stereo mammogram pairs (top) with different scan-lines highlighted in color. The corresponding 1-D luminance cross sectional profiles are shown in the bottom.

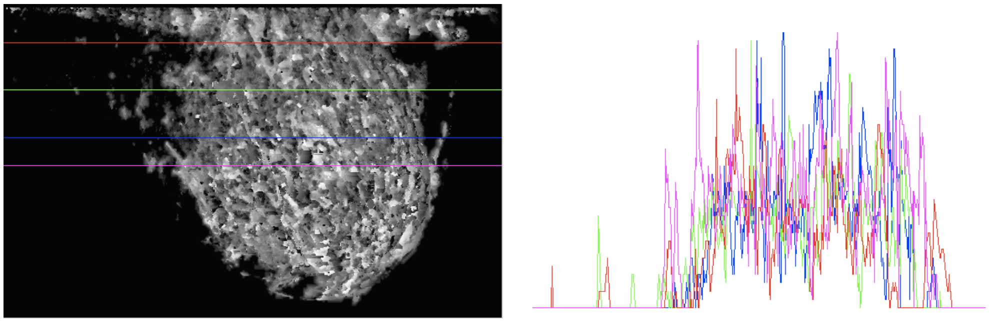

Figure 7.2: Example of a disparity map computed using the proposed model for the stereo pair shown in Fig. 7.3 with different scan-lines highlighted in color (left). The corresponding 1-D disparity cross sectional profiles are shown in the right.

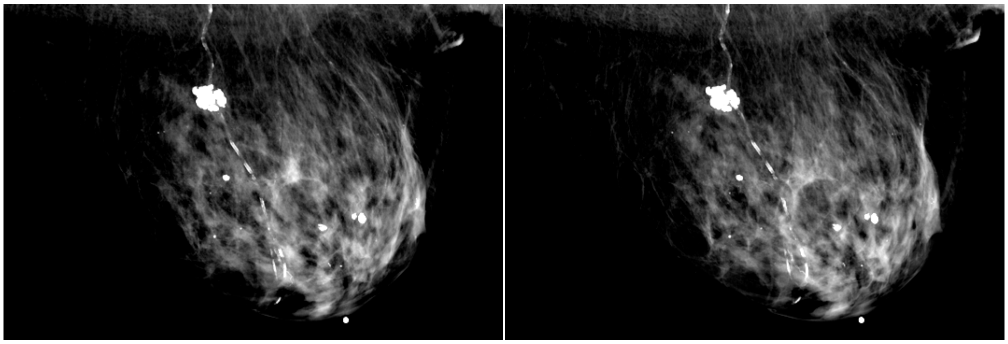

Figure 7.3: A pair of stereo mammogram images.

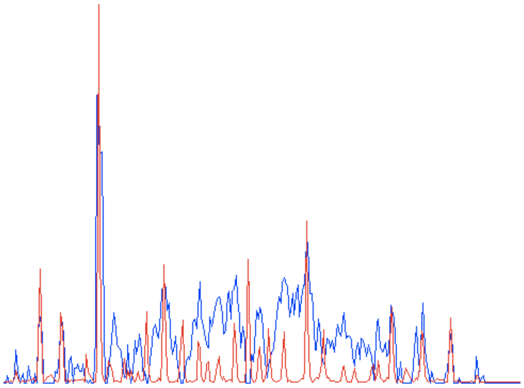

Figure 7.4: Response of the smoothed singularity impulse index (red) when applied to a 1-D luminance cross-sectional profile from a stereo mammogram image (blue).

dense (radio-opaque) tissues, while responding minimally to impulses. We employ both the impulse and edge singularity indices (ψ^1_σ and ψ^2_σ, respectively) to specifically enhance the difficult disparity estimation of critical curvilinear structures such as spicules and vasculature seen on a pair of stereo mammogram images. Their purpose is two-fold: 1) the 2-D impulse and edge singularity indices applied to the reference view of a stereo mammogram pair provide valuable cues to the locations where the disparity smoothness constraint needs to be de-emphasized, and 2) the 1-D impulse index is used to promote curvilinear masses at locations where the 2-D impulse index produced a strong response when applied to the reference view of the stereo mammogram pair. The complete stereo model is defined as follows:

$$\begin{aligned} E(D(\mathbf{x})) = E_P(D(\mathbf{x}))+ \\ (1-w_1(\mathbf{x}))(1-w_2(\mathbf{x}))E_S(D(\mathbf{x}))+ \\ w_2(\mathbf{x})E_C(D(\mathbf{x})), \end{aligned} \tag{7.4}$$

where the term $E_P(D(\mathbf{x}))$ enforces the photometric constraint and is defined as in well-known canonical stereo models, the term $E_S(D(\mathbf{x}))$ enforces the disparity smoothness constraint, and the term $E_C(D(\mathbf{x}))$ promotes curvilinear masses in disparity space. Further, $w_1(\mathbf{x})$ and $w_2(\mathbf{x})$ serve to control $E_S(D(\mathbf{x}))$ and $E_C(D(\mathbf{x}))$, and are defined as functions of the edge index (ψ^2_σ) and the impulse index (ψ^1_σ) applied to the reference image, respectively: $w_1(\mathbf{x}) = 1 - L(\psi^2_\sigma[I_1(\mathbf{x})])$ and $w_2(\mathbf{x}) = 1 - L(\psi^1_\sigma[I_1(\mathbf{x})])$. The function $L(\nu) = exp(-\frac{\nu}{C_T})$ maps the singularity index response to the range $[0, 1]$, where the constant

$C_T \in R$ is a threshold on the index value. $w_1(\mathbf{x}) \to 1$ implies the presence of an edge at the location $\mathbf{x}$ in the reference image, while $w_2(\mathbf{x}) \to 1$ implies the presence of an impulse at the location $\mathbf{x}$. The disparity smoothness constraint E_S is disabled at locations where either $w_1(\mathbf{x}) \to 1$ or $w_2(\mathbf{x}) \to 1$. Likewise, curvilinear masses are promoted via the term E_C in disparity space at locations where $w_2(\mathbf{x}) \to 1$. The disparity smoothness term is defined as $E_S(D(\mathbf{x})) = \lambda_1 \sum_{\mathbf{p} \in N} (1 - w_1(\mathbf{p}))|D(\mathbf{x}) - D(\mathbf{p})|$, where $\mathbf{p}$ is a pixel location in the neighborhood N of $\mathbf{x}$, and λ_1 is a weight indicating the relative importance of the term. We next describe the term $E_C(D(\mathbf{x}))$.

7.3.2 Promoting Curvilinear Masses

We hypothesize that critical, fine-scale curvilinear masses such as vasculature and spicules that are contained in the breast are characterized by large second derivatives and small first derivatives in the direction orthogonal to the axes of the curvilinear masses in disparity space. On the other hand, it has been perviously shown [50] that disparity varies smoothly along image contours projected from 3-D contours. The term E_C embodies both these properties of curvilinear masses and is defined as follows:

$$\begin{aligned} E_C(D(\mathbf{x})) = \lambda_2\{(&|D(\mathbf{q}) - D(\mathbf{x})| + |D(\mathbf{x}) - D(\mathbf{p})|) + \\ &(|D(\mathbf{q}) - 2 * D(\mathbf{x}) + D(\mathbf{p})|)\} - \\ &\lambda_3(\sigma^2 \psi_\sigma^1[D(\mathbf{x})]), \end{aligned} \tag{7.5}$$

where $\mathbf{p}$ and $\mathbf{q}$ are locations in a 3×3 neighborhood of $\mathbf{x}$, and λ_2 and λ_3 represent the relative weights of the individual terms.

We employ the 1-D scale-normalized smoothed singularity index (7.2) in (7.5) to promote large second derivatives and small first derivatives in the direction orthogonal to a curvilinear mass in disparity space. Since (7.5) is applied point-wise at locations of isolated impulses and curvilinear masses estimated by the weighting function w_2 from the reference image of the stereo pair (i.e. the location of the impulses are pre-computed), the factor a in (7.2) is set to 1. Further, the convolution operator in (7.2) is replaced by the simple inner product.

The first two terms in (7.5) augment the singularity index by ensuring that disparity varies smoothly and continuously along the axes of curvilinear masses [50]. Hence, neighboring locations $\mathbf{p}$ and $\mathbf{q}$ are selected to lie along the curvilinear mass, i.e. along a direction orthogonal to the dominant orientation $\theta_D(\mathbf{x})$, whose estimate is obtained from the 2-D singularity index applied to the reference image of the stereo pair. In (7.5), the singularity index ψ_σ is applied to a 1-D sequence of disparity values in a neighborhood of $\mathbf{x}$ lying along the dominant orientation. The size of this neighborhood is dependent on the scale σ of the 1-D gaussian filter in (7.2). The dominant orientation θ_D at each pixel location $\mathbf{x}$ is quantized to lie along one of four directions (east-west, northeast-southwest, north-south, northwest-southeast) surrounding the pixel location $\mathbf{x}$. The locations $\mathbf{p}$ and $\mathbf{q}$ are then chosen to be along a direction orthogonal to the dominant orientation. This is illustrated in Fig. 7.5.

Algorithm 1 Optimization of the Proposed Stereo Model

```
Inputs: Stereo image pair $I_1$ and $I_2$ with $I_1$ taken as the reference image, an
    initial estimate for the disparity map $D$, and the weighting functions $w_1$
    and $w_2$ computed by applying the 2-D singularity index $\psi_\sigma$ on $I_1$.
Output: The final disparity map $D$ for $I_1$.
Set a start temperature $T$.
while $T >$ end temperature do
    for each pixel location $\mathbf{x}$ in $I_1$ do
        $D_1 \leftarrow D(\mathbf{x})$
        Compute $E(D_1) = E_P(D_1) + (1 - w_1(\mathbf{x}))(1 - w_2(\mathbf{x}))E_S(D_1) +$
        $w_2(\mathbf{x})E_C(D_1)$.
        Sample a new disparity $D_2$ uniformly from the range $[D_L, D_U]$.
        Compute $E(D_2) = E_P(D_2) + (1 - w_1(\mathbf{x}))(1 - w_2(\mathbf{x}))E_S(D_2) +$
        $w_2(\mathbf{x})E_C(D_2)$.
        $\Delta E \leftarrow E(D_2) - E(D_1)$
        if $\Delta E < 0$ then
            $D(\mathbf{x}) \leftarrow D_2$
        else
            Select a random number $r$ uniformly from the range $[0, 1]$.
            if $r \leq exp(-\frac{\Delta E}{T})$ then
                $D(\mathbf{x}) \leftarrow D_2$
            end if
        end if
    end for
    $T \leftarrow T - \Delta T$
end while
```

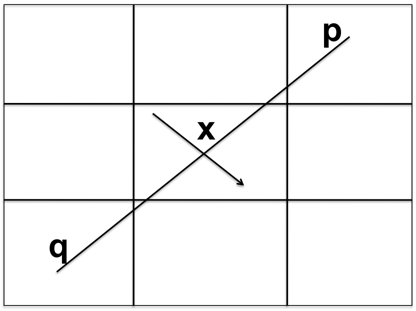

Figure 7.5: The dominant orientation of a curvilinear structure at location **x** is illustrated by the arrow. The locations **p** and **q** are the neighboring locations orthogonal to the dominant orientation.

7.4 Multi Scale Optimization

We employ the Metropolis simulated annealing algorithm to optimize both the proposed stereo model (7.4) and the canonical stereo model (7.1). While slow, the Metropolis simulated annealing algorithm has been shown to produce good disparity estimates, provided the annealing schedule is properly selected [6]. Recently, optimization techniques based on graph cuts [15], [106] and belief propagation [101] have been shown to perform better than simulated annealing for solving global stereo correspondence problems. However, these come with the caveat that simple first or second order priors be used to regularize the disparity, since otherwise the construction of the graph and inference become intractable. While the canonical model is comprised of a simple first order prior, the disparity regularizers in our proposed model are comprised of non-linear higher-order terms which are regulated by the singularity index.

We adopt a hierarchical, coarse-to-fine multi-scale optimization approach that iteratively smooths and sub-samples the stereo images by a factor of two. Simulated annealing is run at each scale and the disparity obtained from a coarser scale serves as the initialization at the subsequent finer scale. The disparity estimated at a coarser scale is linearly interpolated and multiplied by a factor of 2 prior to using it as an initial estimate at each subsequent finer scale. Algorithm 1 gives the pseudo-code for the optimization of the proposed model at a given hierarchical scale. Finally, note that though we do not handle occlusion explicitly in the stereo models, points of half-occlusion could be detected by employing the simple left-right consistency test as a post-processing step on disparities estimated from the left and right stereo images [26] using simulated annealing.

7.5 Experimental Methodology

7.5.1 Creation of Synthetic Mammograms

Real stereo mammogram images lack ground truth disparity data. Hence, only qualitative results of the proposed and canonical stereo models can be shown on real stereo mammograms. In order to quantitatively compare the proposed and canonical stereo models, synthetic stereo images with ground truth disparity data were generated. The synthetic images were created with an $1/f^{\beta}$ background texture. For each stereo pair, the exponent β was a random number uniformly generated on $[1.0, 2.5]$ to resemble mammographic texture [18]. The background texture was assigned a global disparity of 3 pixels.

Five circular regions whose centers and radii were randomly selected in

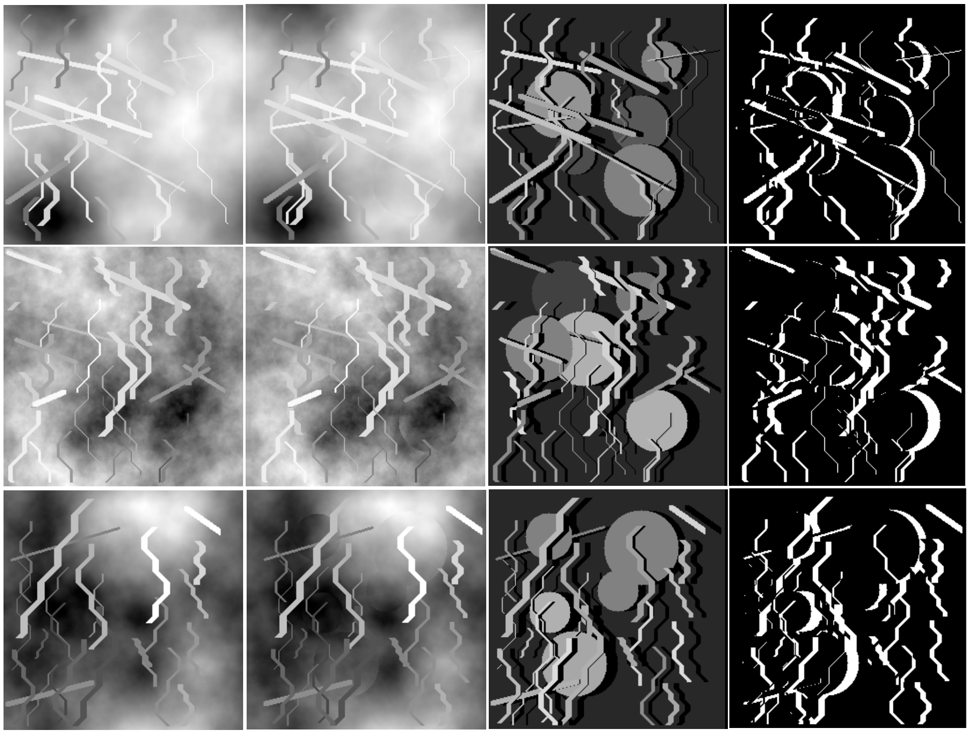

Figure 7.6: Examples of synthetic stereo images. The reference image is shown in column 1, while the ground truth disparity and the ground truth occlusion maps are shown in columns 3 and 4, respectively. Occluded pixels in the ground truth disparity map have 0 intensity.

the reference image were horizontally displaced by a random disparity value uniformly selected from the range 5-16 pixels in the other image of the stereo pair. Each center coordinate was constrained to lie within the image plane and to be the only coordinate within a 25×25 window around it. The radii of the circular regions were uniformly selected from the range 20-40 pixels.

Curvilinear structures intended to mimic those found on real mammograms were then overlaid on the background texture in the reference image. The intensity along the curvilinear structures was set to a constant whose value was computed as the sum of the background intensity at the first coordinate of the curvilinear structure and a random positive offset. The number, orientation, starting coordinate, length, width, and the disparity of the curvilinear structures in each stereo pair were randomly selected. The number of curvilinear structures was uniformly generated on $[20, 30]$, each starting coordinate was constrained to lie within the image plane and to be the only coordinate within a 25×25 window around it, while the length, width, and disparity of each curvilinear structure was uniformly selected from 30-150 pixels, 1-8 pixels, and 11-21 pixels, respectively.

Each curvilinear structure was created as follows: a random number was uniformly generated on $[0, 1]$. If the value of this random number was greater than 0.75, then the curvilinear structure was created as a straight line between two end points using the Bresenham line algorithm [16]. Given the start coordinate, the end coordinate was determined based on the length and orientation of the straight line. The orientation was a random number uniformly selected from 5-30 degrees. If the random number was less than 0.75, the curvilinear structures were generated using a contour traversal algorithm with tortuousity. First, given the start coordinate, a direction of growth was randomly selected. This could either be south (vertical), southeast (diagonal right), or southwest (diagonal left). After traversing in this direction for $1/9^{th}$ of the length, the direction of traversal was changed. If the previous direction of traversal was either southeast or southwest, then the new direction of traversal could only be south, while if the previous direction was south, then the new direction was randomly selected to be either southeast or southwest. The traversal continued till the pre-specified length of the contour was reached or the contour grew out of the image plane. Fig. 7.6 illustrates three examples of synthetic stereo images with ground truth disparity data and occlusion maps created in this way.

7.5.2 Parameter Setting

Here, we describe the parameter settings that are used in both the canonical and proposed stereo models. The parameter λ_1, which controls the disparity smoothness constraint is common to both the models. Typically, λ_1 is set according to the application. For example, if the 3-D scene is composed of primarily piece-wise smooth surfaces, then λ_1 is set to a large value to enforce the disparity smoothness constraint tightly. On the other hand, in the stereo-mammography application, the breast is composed of very few piece-wise smooth surfaces, that usually arise from solid breast masses and dense (radio-opaque) tissues. Hence, we set $\lambda_1 = 1$ in both the models. If λ_1 is set to a large value, then the resulting disparity space will be over-smoothed. We used an 8-connected square neighborhood ($N = 8$) to evaluate the smoothness constraint.

The parameters λ_2 and λ_3 are specific to the proposed model and control the continuity and promotion of curvilinear masses in disparity space, respectively. We set $\lambda_2 > \lambda_3$, thereby placing a stronger emphasis on disparity continuity along the direction of curvilinear masses. We set $\lambda_2 = 100$ and $\lambda_3 = 5$. These parameters were found to yield good performance on the synthetic stereo pairs. Note that one could also adopt the more computationally intensive grid-search and cross-validation processes [56] to arrive at values for the parameters λ_2 and λ_3.

The number of hierarchical scales used in the multi-scale optimization process was set to 3. At each hierarchical scale of optimization, both the 2-D impulse and edge singularity indices were applied to the reference view of the stereo pair at that scale. The 2-D impulse index was applied to detect fine scale curvilinear structures at each hierarchical scale of optimization, and hence we set $\sigma = 1.5$ pixels. Similarly, the scale σ of the 1-D impulse index that is used to promote curvilinear masses in disparity space was set to 1.5 pixels. The 2-D edge index was computed over 5 scales at each hierarchical scale of optimization, where the lowest scale was 1.5 pixels, and each subsequent coarser scale was larger than the previous finer scale by a factor of $\sqrt{2}$.

A linear annealing schedule comprised of 200 iterations was employed with a start temperature of 10.0 and end temperature of 0.01 for both the proposed and canonical stereo models. The temperature decrement ΔT (see Algorithm 1) was set to 0.05, while the integer disparity range ($[D_L, D_U]$ in Algorithm 1 was set to $[1, 21]$ pixels. The initial disparity estimate at the coarsest scale was set to 1 pixel everywhere.

7.5.3 Performance Measure

The stereo models were evaluated using the percent bad pixel error measure $B = \frac{1}{P}\sum\limits_{(x,y)}(|D(x,y) - GT(x,y)| > \delta_D)$, where D is the estimated disparity, GT is the ground truth disparity, P is the total number of pixels in the image, and δ_D is the disparity error tolerance. We set $\delta_D = 1$ pixel. We report the following percent bad pixel error measures: a) percent bad pixel error measure along curvilinear structures (B_c), b) overall percent bad pixel error measure (B), c) percent bad pixel error measure evaluated only for the non-occluded curvilinear pixels (B_{cnocc}), and d) overall percent bad pixel error measure evaluated only for non-occluded pixels (B_{nocc}). Note that B_{cnocc} and B_{nocc} were computed using the occlusion maps created from the ground truth disparity data. The difference in the performance measures of the canonical and proposed models were evaluated for statistical significance using the Wilcoxon Sign Rank (WSR) test [91].

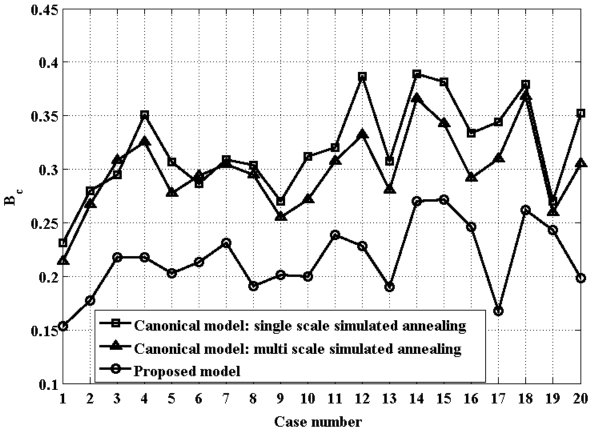

Figure 7.7: Percent bad pixel error of the proposed and the canonical single-scale and multi-scale stereo models along the curvilinear structures.

7.5.4 Real Mammograms

The proposed and the canonical stereo models were also run on a set of five real stereo mammogram images (courtesy Emory University, Atlanta, GA). Figs. 7.3 and 7.14 shows three example stereo mammograms. For computational efficiency, the images were sub-sampled by a factor of 4 prior to running both stereo models. The algorithm settings used for the real stereo mammogram images were identical to the synthetic stereo images, except the disparity range was set to $[1, 35]$ pixels. This range was arrived at by manually inspecting many landmark points that were visible on both the images of the stereo mammogram pairs.

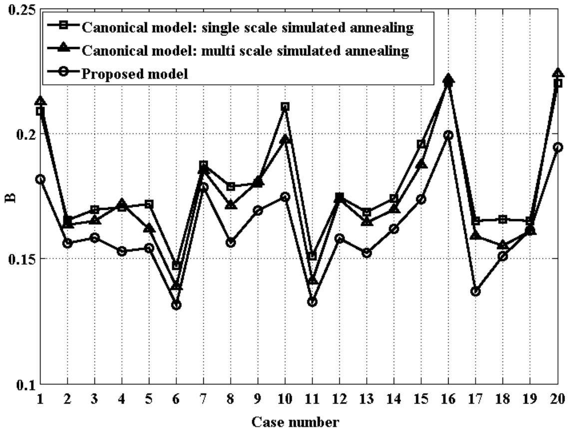

Figure 7.8: Overall percent bad pixel error of the proposed and the canonical single-scale and multi-scale stereo models.

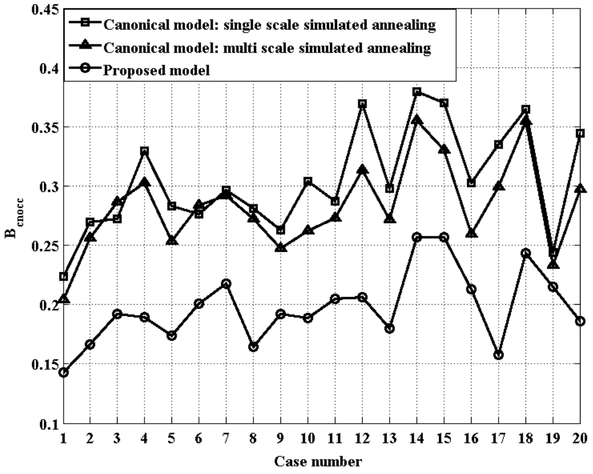

Figure 7.9: Percent bad pixel error of the proposed and the canonical single-scale and multi-scale stereo models along the non-occluded curvilinear pixels.

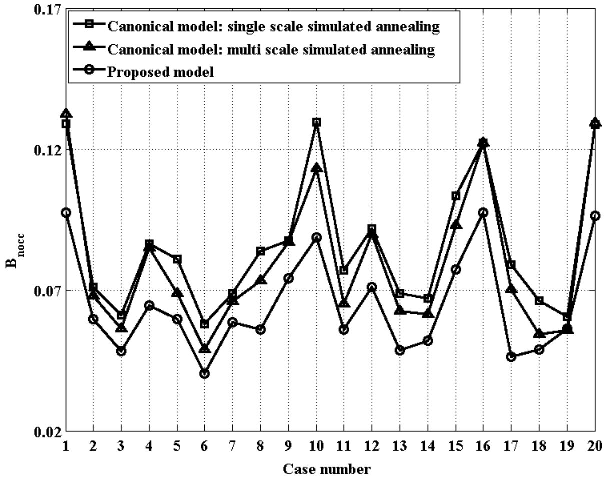

Figure 7.10: Overall percent bad pixel error of the proposed and the canonical single-scale and multi-scale stereo models for only the non-occluded pixels.

7.6 Results

Fig. 7.7 plots B_c computed for the proposed and the canonical stereo models for the 20 synthetic stereo images. As can be seen from Fig. 7.7, the proposed model clearly outperforms the single-scale and multi-scale canonical models in estimating disparity along the curvilinear structures. The difference in performance is statistically significant as evaluated via the WSR test (*p-value* < 0.0001 for both the single-scale and multi-scale canonical models). Also, the performance of the multi-scale canonical model is better than the single-scale model. Fig. 7.8 plots B computed for the proposed and the canonical stereo models for the 20 synthetic stereo images. Again, the proposed model performs better than the single-scale and multi-scale canonical models. The difference in performance is again statistically significant (WSR *p-value* $<$

0.001 for both the canonical models). Figs. 7.9 and 7.10 plot B_{cnocc} and B_{nocc}, respectively. The better performance of the proposed model is evident from these figures. Again, the difference in performance is statistically significant.

Fig. 7.11 shows the ground truth disparity (column 1) and the estimated disparity maps from the proposed model (column 2) and the multi-scale canonical model (columns 3) for the 3 synthetic stereo images illustrated in Fig. 7.6. Fig. 7.12 shows the terms w_1 and w_2, which are used in the proposed model. The disparity maps shown in Fig. 7.11 are pre-multiplied by the ground truth occlusion maps (see Fig. 7.6). Hence, occluded pixels have 0 intensity (black) in the disparity maps in Fig. 7.11. The red circles superimposed on the disparity map of the canonical model illustrate some of the key regions where the canonical model produces erroneous disparity estimates and the proposed model does not. The disparity along the curvilinear masses in these regions is affected by the background disparity in the canonical model, resulting in blurring, while in the proposed model, the disparity is better estimated due to the de-emphasis of smoothness at impulse and edge locations and explicit promotion of curvilinear masses in the disparity space.

We also investigated the impact of the weighting function w_2 in estimating disparity along the curvilinear masses in the proposed model. Fig. 7.13 plots B_{cnocc} for the stereo models with and without the term w_2. Fig. 7.13 suggests that the term w_2 helps produce better disparity estimates along the curvilinear structures. The difference between the two models (with and without w_2) was statistically significant (WSR *p-value* < 0.05).

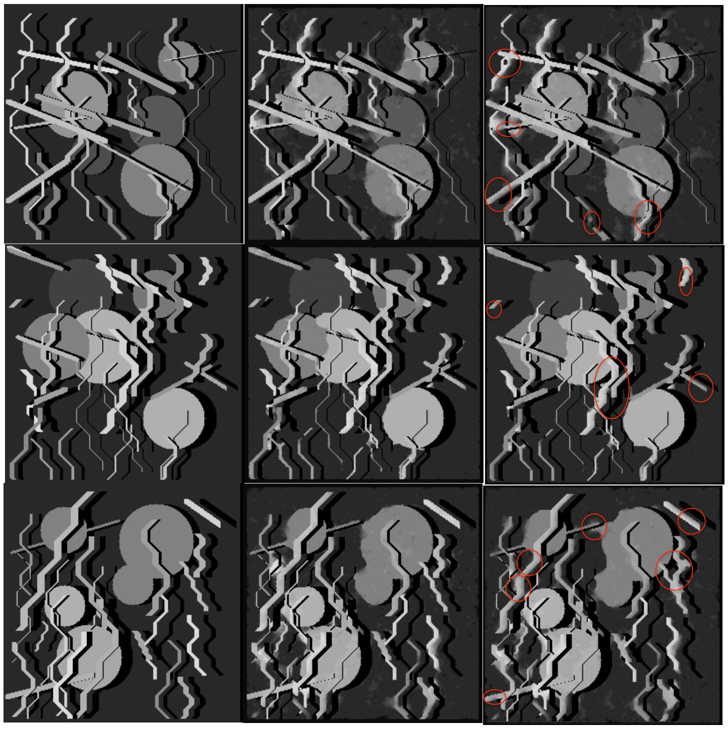

Figure 7.11: Results of the proposed and the multi-scale canonical stereo models on the synthetic stereo images illustrated in Fig. 7.6. Column 1: ground truth disparity, Column 2: disparity estimated from the proposed model, and Column 3: disparity estimated from the canonical model. The red circles superimposed on column 3 show the key regions were the disparity estimates produced by the canonical model were incorrect.

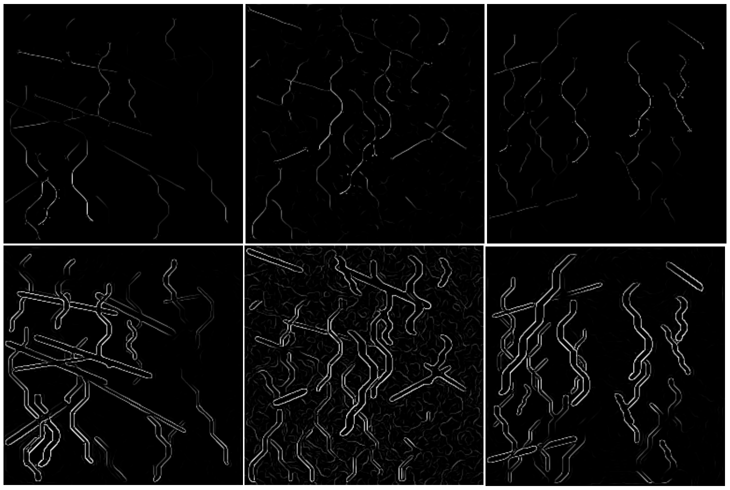

Figure 7.12: The weighting functions w_2 (top row) and w_1 (bottom row) computed for the synthetic stereo images illustrated in Fig. 7.6.

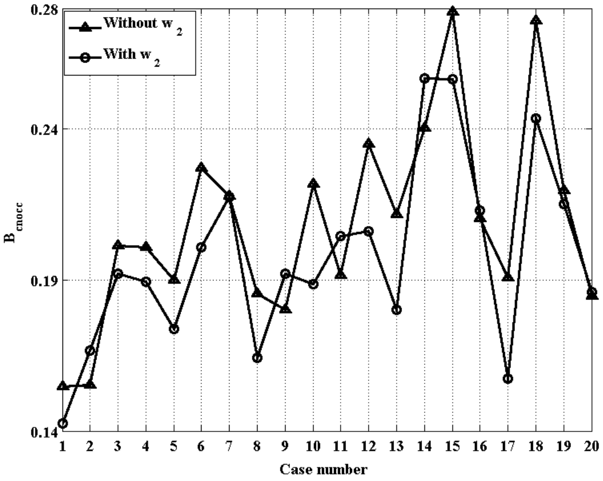

Figure 7.13: Percent bad pixel error with and without the term w_2 along the non-occluded curvilinear pixels.

Fig. 7.15 shows the disparity maps estimated by the proposed and the canonical stereo models for the three stereo mammograms shown in Figs. 7.3 and 7.14. Fig. 7.16 shows the terms w_1 and w_2 computed from the reference image of the stereo pair, which are used in the proposed model. In order to visualize the results of the two models better, Fig. 7.17 shows the close up of 4 representative disparity regions comprising curvilinear structures. Fig. 7.17 reveals blurring of the curvilinear structures in the regions indicated by the red circles in the disparity map produced by the canonical stereo model (right column), while the same appear sharper in the proposed model. While it is hard to ascertain these results quantitatively due to lack of ground truth disparity data, the proposed stereo model can be observed to preserve the curvilinear masses in the disparity space better than the canonical stereo model.

7.7 Summary

We have created a new stereo correspondence model for estimating disparity for a given pair of stereo mammogram images. The algorithm employs the recently developed singularity index described in chapter 5 to reliably detect locations of impulse and edge singularities in the reference image of the stereo pair. The singularity index is used to de-emphasize smoothness and promote curvilinear structures in the disparity space. Extensive experimental results on synthetic and real data show that the proposed model offers superior performance over the canonical stereo model.

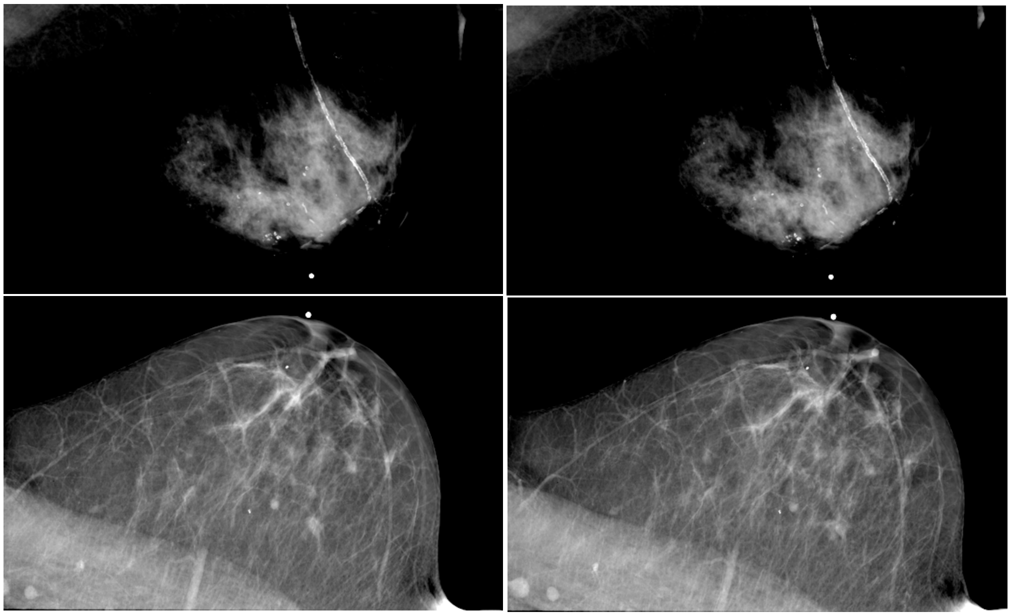

Figure 7.14: Stereo mammograms courtesy Emory University, Atlanta, GA. The reference image is shown in the left column.

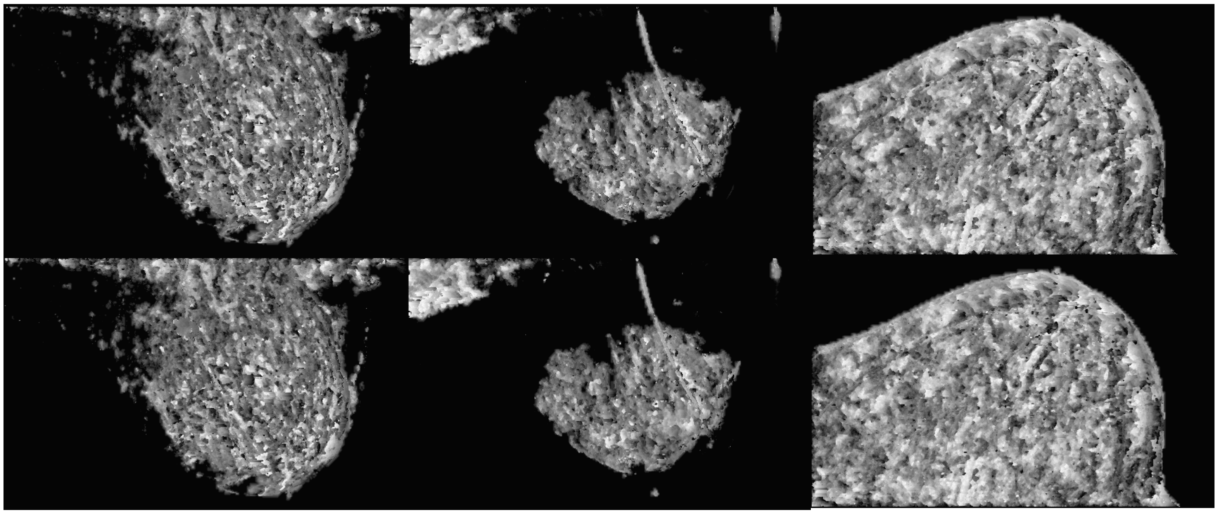

Figure 7.15: Results of the proposed stereo model (top row) and the canonical stereo model (bottom row) on the stereo mammograms illustrated in Figs. 4.1 and 7.14.

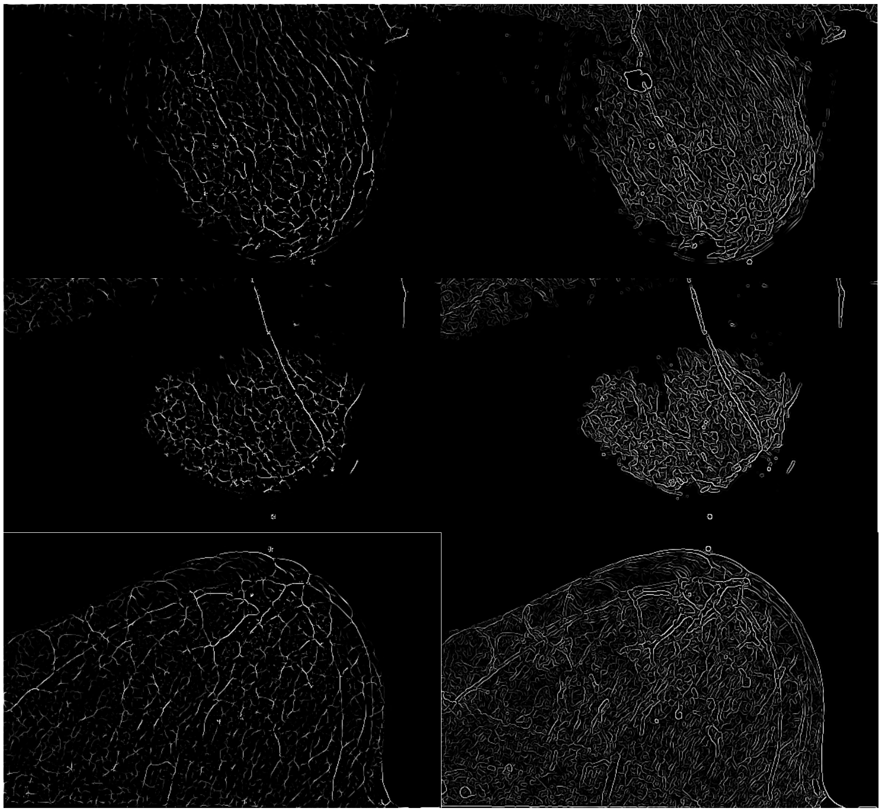

Figure 7.16: The weighting functions w_2 (left) and w_1 (right) computed for the reference image of the stereo mammogram pairs (left column) illustrated in Figs. 4.1 and 7.14.

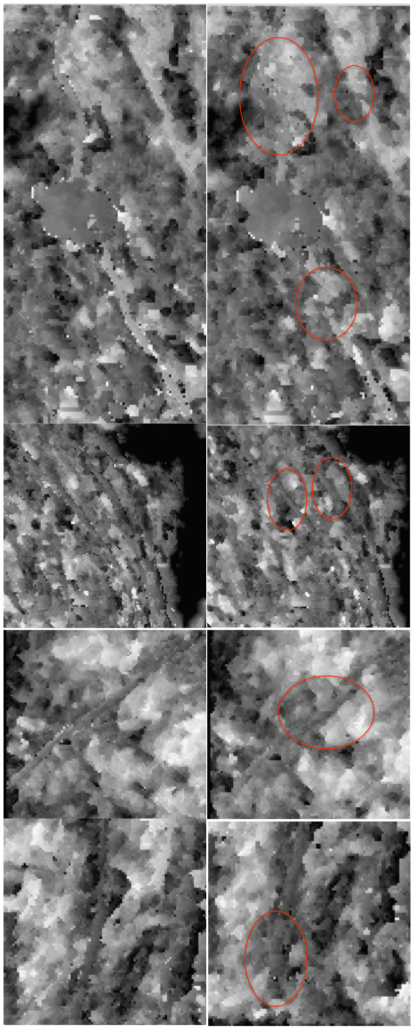

Figure 7.17: Close up of the estimated disparity maps with the proposed stereo model in the left column and the corresponding region from the canonical model in the right column. The red circles superimposed in the right column indicate regions of blurring in the disparity map produced by the canonical model.

Chapter 8

Stereoscopic Viewing of Breast Tomosynthesis Projection Images

8.1 Introduction

Stereo viewing is a powerful technique applicable to not only stereo mammography, but other new 3-D breast imaging modalities such as breast tomosynthesis and breast CT. In breast tomosynthesis imaging, 15-30 x-ray projections of the breast are typically acquired over a limited angular range of 15-50 degrees [76], [24]. Multiple slices of the breast can be synthesized from the projection images and these slices provide 3-D visual information about the structures within the breast. Yet, due to limited angular sampling, tomosynthesis is often regarded as a quasi-3-D modality as the resolution between slices (in the direction perpendicular to the detector plane) is poorer than the in-plane (plane parallel to the detector plane) resolution [45]. The typical in-plane resolution provided by present-day tomosynthesis systems is in the range of 100-150 micrometers per pixel while the resolution between slices is around 1 millimeter.

The imaging geometry of breast tomosynthesis naturally lends itself to stereo visualization, provided that the angular separation between the projection images is carefully selected. Stereo mammography has been shown to be effective in improving the sensitivity of breast cancer detection and reducing unnecessary patient recalls [31]. However, in the clinical studies conducted to date, the dose of the stereo mammographic examination was twice the dose of a regular screening mammographic examination [31]. The x-ray dose under which each tomosynthesis projection image is acquired is much lower than the dose under which a standard mammogram is acquired. Consequently, tomosynthesis projection images have poor contrast due to a low signal to noise ratio [82] when compared to traditional x-ray mammograms. However, in a recently published study, Webb et al. [103] showed that radiologists performed significantly better in detecting simulated breast masses on tomosynthesis projection images viewed on a stereo display than on regular mammograms. In this chapter, we present an observer study that we conducted with experienced breast imaging radiologists in which we assessed the feasibility of stereo viewing of low-dose tomosynthesis projection images depicting real breast masses. We begin with a description of the main motivation behind this study.

8.2 Motivation

Clinical studies on breast tomosynthesis have shown the modality to be more advantageous than conventional mammography for the detection of non-calcified lesions, particularly in dense breasts [85], [86]. Breast tomosynthesis has also shown to reduce patient recalls and enable higher sensitivity when used as an interpretative aid along with digital mammography. Poplack *et al.* [81] reported that the image quality of tomosynthesis was equivalent to or superior to diagnostic mammography in the majority of cases and that the recall rate can be substantially reduced when tomosynthesis is used as an adjunct modality to digital screening mammography. Gur *et al.* [36] showed that the combined use of breast tomosynthesis and full field digital mammography can result in a significant reduction in the recall rate as compared to digital mammography alone. In another non-blinded consensus study, Andersson *et al.* [2] determined that cancers were often more visible on breast tomosynthesis images as compared to one-view and two-view mammograms, which suggest that using breast tomosynthesis might lead to an increase in sensitivity. These studies strongly suggest that 3-D breast tomosynthesis provides more visual information in the form of 3-D structural information within the breast to the radiologist than does conventional mammography. Breast tomosynthesis was also recently approved by the U. S. Food and Drug Administration (FDA) for clinical use.

In the clinical studies that have been conducted on breast tomosynthesis so far (e.g., [36], [81], [2]), the imaging data has been interpreted using either the raw projection images (similar to digital mammography interpretation) or the reconstructed slices (using either a slice-by-slice or cine mode). Yet, there is still uncertainty as to which reading mode is the best for reliable and efficient interpretation of breast tomosynthesis data. For instance, Good *et al.* [33] conducted a pilot observer study to assess the ergonomic and diagnostic performance-related issues associated with interpretation of breast tomosynthesis examinations. While several radiologists did perceive breast tomosynthesis examinations to be better than the corresponding full field digital mammography examinations, the authors suggested that much work needed to be done in order to understand the optimal reading mode for breast tomosynthesis images if the modality were to be routinely used in clinical practice. A key point is that the radiologist is faced with interpreting as many as 300 slices. The average time required for interpreting tomosynthesis examinations alone or in conjunction with full field digital mammography is higher than that of reviewing full field digital mammography examinations alone. Thus, a vast amount of data will be generated if tomosynthesis is used for screening, given the large number of screening examinations. Visual analysis of 3-D tomosynthesis images may prove to be even more repetitive and fatiguing than traditional mammography. Hence, there is a need for efficient reading strategies for interpreting breast tomosynthesis data.

Unlike the current cine or slice-by-slice viewing modes, stereo visualization of tomosynthesis projections has the potential to reveal the 3-D structure of the breast. In this study, we evaluated two important questions: 1) Can radiologists perceive 3-D information when viewing the low dose projection images stereoscopically? and 2) Can mass detection be reliably performed on tomosynthesis projection images using stereoscopic viewing?. These questions were geared towards finding whether stereo viewing of tomosynthesis projection images could be a viable reading mode for interpreting breast tomosynthesis data in the future.

8.3 Material and Methods

8.3.1 Data Set

Breast tomosynthesis images were provided by Hologic, Inc. (Bedford, MA, USA). A total of 47 craniocaudal breast tomosynthesis cases, which were comprised of 23 biopsy-proven malignant masses and 24 normals, were used in this study. Each case consisted of 15 projection images, which spanned an angular range of approximately 15 degrees. A stereo pair was formed by selecting two images that were approximately ± 4 degrees apart from the zero angle projection (total separation of 8 degrees). No other criterion was used in selecting the projection images of the stereo pair. Previous studies with stereo mammography have shown that in order to achieve good depth perception of breast tissue, the angle of separation between the two images of the stereo pair should be between 6 and 10 degrees [31].

An experienced breast imager with close to 20 years of experience in interpreting screening mammograms certified each case that was used in this study as being a normal or a lesion case. This breast imager, to whom we will refer as the truth radiologist, determined the ground truth location information on each stereo pair of projection images by visually comparing the lesion depicted on the projection images and the corresponding reconstructed slices. For the reconstructed slices, Hologic, Inc. provided the ground truth location information, and the truth radiologist used this information to determine the ground truth location on the stereo pair of projection images. The truth radiologist also assessed the tissue density, lesion subtlety, and certified each stereo

Fatty	Scattered densities	Heterogeneously dense	Extremely dense
10	21	13	3

Table 8.1: Number of cases according to breast density category in the dataset.

pair as having sufficient contrast for interpretation. Table 8.1 illustrates the distribution of densities represented in the cohort of cases used for the study.

8.3.2 Preprocessing

We manually preprocessed the stereo pairs of the raw projection images to enhance their contrast. This was achieved by manually adjusting the DICOM window width/window level parameter to obtain satisfactory contrast in the two images of the stereo pair. The DICOM window width/window level parameter was set to the same value in both images of the stereo pair. Subsequently, we corrected the histograms of the two images by shifting the median of one histogram with respect to the other such that they were identically matched in shape. To perceive stereoscopic effect, the images of the stereo pair had to be rotated by 90 degrees in the clockwise direction such that the left breast was displayed as though the patient was in a prone position, while the right breast was displayed as though the patient was in a supine position. This display of images, while different from the display of conventional mammographic images, provides horizontal parallax, which in turn induces stereoscopic perception. The stereo display mode used in our study was similar to that used by Getty et al. in their stereo mammography study [31]. Finally, a white dot was randomly placed on either the left or the

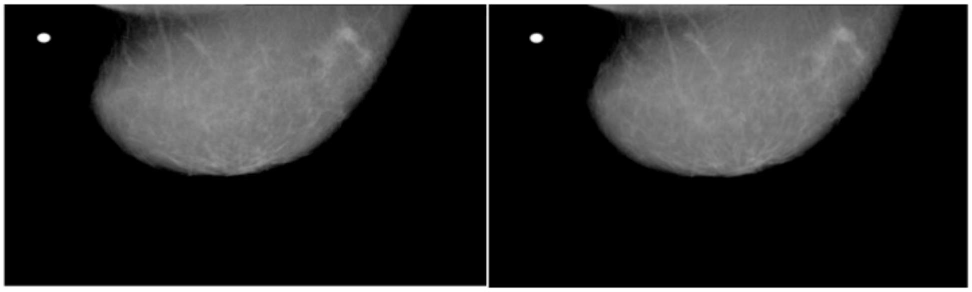

Figure 8.1: Processed stereo pair of tomosynthesis projection images of the left breast.

right side of the non-tissue region of the images. The white dot was used as a reference by readers while commenting on the location of the abnormality found on the image. Figure 8.1 illustrates an example of a processed stereo pair of tomosynthesis projection images of the left breast used in our study.

8.3.3 Study Setting

The stereo pairs were displayed on a Planar PL2010M stereo display (maximum resolution 1600x1200) manufactured by Planar Systems, Inc. (Beaverton, OR, USA). This was not a standard mammographic display, but a research workstation. The Stereoscopic PlayerTM (3dtv.at, Linz, Austria) software was used to load the stereo pair on the stereo display system. A stereo pair loaded on this stereo display system can be fused by a human having normal stereo acuity by wearing lightweight passive cross-polarized glasses.

The study was blinded and was held over two phases. The main objective of the first phase was to assess if the radiologists could perceive 3-D information from the low dose projection images. The main readers who took part in the first phase of the study were two experienced breast imagers with a collective experience of more than 30 years in interpreting screening mammograms. In addition to these two breast imagers, a radiologist who specialized in reading computed tomography (CT) scans, and an imaging physicist participated in the study design as trial readers. The trial readers only evaluated the study design.

A questionnaire was devised with the help of the truth radiologist to collect information on how well the study readers perceived 3-D information on a scale of 1-3 (1-no 3-D perception, 2-moderate 3-D perception, and 3-excellent 3-D perception). The questionnaire also contained specific questions for describing the morphological characteristics of the most suspicious finding identified by the readers. The morphological characteristics of the lesions were described using the American College of Radiology (ACR) Breast Imaging and Data System (BI-RADS) descriptors [25]. The readers were asked to describe the lesion subtlety on a scale of 1-5, where a rating of 1 meant that the lesion was extremely subtle, and a rating of 5 meant that the lesion was extremely obvious. The readers were also asked to characterize the parenchymal pattern according to one of four BI-RADS categories: mainly fatty, scattered fibroglandular densities, heterogeneously dense, and extremely dense [25].

The first phase of the study was split across two sessions, which were held under the same ambient lighting conditions. In the first session, 24 cases were shown to the readers, while in the second session, 23 cases were shown. Each session had a mix of true positive and true negative cases that were shown in a random order. Each reader was shown three stereo tomosynthesis pairs of the breast (1 mass, 2 normals), which were not a part of the study set in a training phase at the start of the first session to get accustomed to the passive cross-polarized glasses, the stereo display, and the questionnaire. A break of at least 30 minutes was scheduled between the two sessions. Each reader was free to adjust the seating and viewing distance in order to best perceive the 3-D view. As soon as a stereo pair was loaded on the stereoscopic display system using the Stereoscopic Player^{TM}software, a stopwatch was started and the readers were given at least 20 seconds to examine the display. This was done to ensure that the readers had adequate time to fuse the stereo pair to form the cyclopean view and become comfortable with their 3-D perception. If in those 20 seconds the readers did not indicate that they were ready to make an interpretation, then they were prompted to see if they were ready. The readers were free to take more time if needed or indicate that they were ready to make an interpretation even before the 20 seconds had elapsed. The images of each case were presented on the display until the readers had answered all the questions for that particular case. Only the craniocaudal stereo pairs were available to the readers, with no additional images or display modes. Further, the readers were not allowed to modify any of the display parameters such as magnification or contrast.

The main objective of the second phase was to assess if mass detection could be reliably performed using stereo viewing of tomosynthesis projection images. Here, we expanded on the first phase of the study and compared the mass detection performance of stereo viewing vs. monoscopic (mono) viewing. Four experienced breast imagers with a collective experience of more than 35 years in interpreting screening mammograms participated in the study as blinded readers. One reader was common to both phases. Like in the first phase, the second phase of the study was also held over two sessions, with a mix of true positive and true negative cases shown in a random order. A total of 23 cases were shown in one session, and 24 cases in the other session. In each session, the reader was shown each case twice, once in the stereo mode, and once in the mono mode in a random order. In the mono mode, the same image of the stereo pair was displayed on both monitors of the stereo display system. The reader was not told what the current viewing mode was. Each case was shown on the display system to the readers till the readers had completed interpreting the case. The cross polarized stereo glasses were used by the readers throughout the session, even while viewing the images monoscopically. The readers were not allowed to modify display parameters such as magnification and contrast.

In the second phase, the readers were asked to provide a binary decision on whether they saw a mass for which they would initiate a diagnostic workup or not, and provide a confidence score in the range of 0-100 that indicated their confidence in the presence of the mass. A confidence score of 0 meant that the readers were 100% certain that there was no mass, while a confidence score of 100 meant that the readers were 100% certain that there was a mass. The binary decisions were collected, as they closely resembled how the readers would operate in an actual clinical setting. The confidence scores, which indicated the readers confidence in the presence of an abnormality, were collected for analyzing the readers performance in an experimental setting using the Receiver Operating Characteristic curve (ROC). Previous works (e.g., [37]) on analyzing binary and continuous/multi-category ratings in mammography observer studies have demonstrated that these two decision-making processes yield similar reader performances, even though the binary true positive fraction (TPF) and false positive fraction (FPF) operating points do not always lie on the ROC curves but in their vicinity.

The Randot stereo acuity test was administered on every reader who participated in this study including the truth radiologist, since stereo perception is an innate ability and not all human beings view stereoscopic images equally well. In fact, 4-10% of humans exhibit some degree of stereo deficiency [88]. All the readers passed the Randot stereo acuity test.

8.3.4 Statistical Analyses

Statistical analyses were conducted to assess how well the study readers perceived 3-D information within the breast and their diagnostic performance under the stereo and mono viewing modes. Inter-reader agreement between the two readers in their descriptions of the shape and margin properties, and their ratings of lesion subtlety and tissue density collected during the first phase were quantified using percent agreement. We did not use the Kappa statistic for assessing inter-reader agreement as we found that the Kappa statistic values indicated a low agreement when the percent agreement was very high. This is a well-documented problem in the statistics literature with the Kappa statistic as the Kappa is very sensitive to trait prevalence in the population under consideration. This problem is commonly referred to as the Kappa paradox [38].

From the binary decisions collected in the second phase, we computed the overall binary TPF and FPF for each reader and each viewing mode. Similarly, the empirical ROC curves were generated using the confidence scores for each reader and each viewing mode. We first analyzed whether the binary decisions and the continuous ratings resulted in similar reader performances. A bootstrap analysis was carried out to ascertain this. This analysis was similar to what was described by Gur *et al.* [37] in their work comparing mammography reader performances under binary and continuous/multi-category ratings. We summarize the key steps:

Let us suppose that the i^{th} reader is denoted by R_i and that for this reader, and a given viewing mode (stereo or mono), the binary TPF and FPF obtained were $(TPF_{R_i}^{bin}\ \ FPF_{R_i}^{bin})$. We then evaluated the TPF at the binary FPF, $FPF_{R_i}^{bin}$, from the corresponding ROC curve. We used linear interpolation for values of $FPF_{R_i}^{bin}$ that were not present in the list of FPF values used to generate the empirical ROC curve. Let us denote the linearly interpolated TPF value from the ROC curve by $TPF_{R_i}^{roc}$. The signed vertical difference $TPF_{R_i}^{bin} - TPF_{R_i}^{roc}$ yields a measure that is indicative of how similar the reader performances are under the binary decision and continuous confidence scores.

To asses whether the signed difference in sensitivities was significant or not, we performed bootstrap sampling. For each reader, we separately re-sampled the mass cases and the normal cases independently to ensure that the final sample had the same number of mass and normal cases as in the original dataset. It is important to note that we did not re-sample cases across readers, and each reader's data was analyzed independently of the others. We generated 5000 bootstrap samples, and for each sample the signed vertical difference between the binary and the ROC TPF values were computed as described above. The mean-subtracted bootstrap difference distribution was then used to evaluate the two-sided bootstrap p-value with the test statistic being the signed vertical difference $TPF_{R_i}^{bin} - TPF_{R_i}^{roc}$ computed from the observed data.

The area under the ROC curve (AUC) and the standard deviation in AUC were also computed from the observed data for the stereo and the mono viewing modes. Further, since breast imaging radiologists usually operate at sensitivities greater than 90%, the partial AUC values for 90% and 95% sensitivities were also computed using the pROC software package [90]. The differences in partial AUC values for 90% and 95% sensitivities were statistically assessed [90].

8.4 Results

The two readers who took part in the first phase of the study perceived moderate or excellent 3-D information in 89.36% (42/47) and 93.62% (44/47) of the cases. Both readers remarked that the 3-D perception was better in cases that depicted rich vasculature and other linear structures. A possible reason behind this could be the fact that the stereo matching process in the human visual system is aided by the presence of strong matching features in the two images of the stereo pair.

For the BI-RADS mass shape, mass margin, and assessment ratings, the two readers agreed on 7 out of 16 (43.75%), 11 out of 16 (68.75%), and 12 out of 16 (75%) cases, respectively. For mass subtlety and tissue density ratings, the two readers agreed on 12 out of 16 (75%) and 35 out of 47 (74.47%) cases, respectively. There were only 16 lesions that were correctly detected by both the readers and hence the denominator while computing the BI-RADS percent agreement is 16 as opposed to 23 (the total number of lesion cases). Additionally, this small sample size also tends to exaggerate the lack of agreement.

Figs. 8.2 and 8.3 illustrate the ROC curves for the four readers along with the binary operating points for each reader for the mono and the stereo viewing modes, respectively. As can be seen from Figs. 8.2 and 8.3, the binary operating points do not necessarily lie on the ROC curves. Table 8.2 lists the bootstrap p-values comparing the reader performance using the binary and continuous rating scales for the mono and the stereo viewing modes. Out of the 8 experimental conditions (4 readers, 2 viewing modes) a statistically significant difference in performance between the binary and continuous ratings was observed for only one condition: the mono viewing mode of reader 2. For the remaining experimental conditions, no significant differences in performance

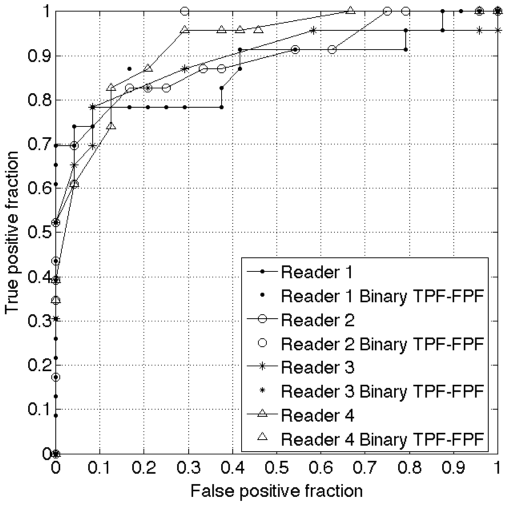

Figure 8.2: ROC curves and the corresponding binary operating points depicting the performance of the four readers under the mono viewing mode.

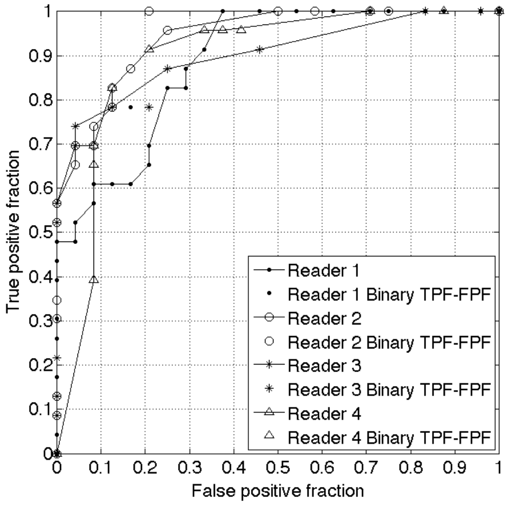

Figure 8.3: ROC curves and the corresponding binary operating points depicting the performance of the four readers under the stereo viewing mode.

Reader No.	Bootstrap p-value (Mono mode)	Bootstrap p-value(Stereo Mode)
1	0.133	0.053
2	0.036	0.169
3	0.466	0.194
4	0.403	0.211

Table 8.2: Bootstrap p-values (two sided) comparing the difference in performance between the binary decision and continuous confidence scores ($p < 0.05$ indicates statistical significance).

were observed between the binary and the continuous confidence scores. Table 8.3 lists the AUC values and their standard deviations, the partial AUC values at 90% and 95% sensitivities, and p-values indicating whether the differences in partial AUC values at 90% and 95% sensitivities were significant or not for the mono and the stereo viewing modes. From Table 8.3, we see that the partial AUC values at 90% and 95% sensitivities were significantly better for readers 1 and 2 under the stereo viewing mode than when compared to the mono viewing mode. For readers 3 and 4, the difference in the partial AUC values at 90% ad 95% sensitivities was not statistically significant.

8.5 Summary

In this study, we have assessed the feasibility of viewing low dose breast tomosynthesis projection images on a stereoscopic display. Our findings suggest that 3-D information is perceived when low dose tomosynthesis projection images are viewed stereoscopically. Further, stereo viewing could help in the interpretation of breast tomosynthesis cases. The results of our pilot study with four readers suggest that stereo viewing might yield equivalent or better detection performance, particularly at sensitivities greater than 90%.

Reader No.	Mono AUC	Stereo AUC	pAUC at 90% and 95%TPF (Mono)	pAUC at 90% and 95%TPF (Stereo)	p-value comparing partial AUC at 90% TPF between mono and stereo viewing modes	p-value comparing partial AUC at 95% TPF between mono and stereo viewing modes
1	0.87 ± 0.05	0.887 ± 0.04	0.022, 0.006	0.064, 0.031	0.037	0.013
2	0.888 ± 0.04	0.942 ± 0.03	0.033, 0.014	0.071, 0.032	0.026	0.037
3	0.889 ± 0.05	0.904 ± 0.04	0.028, 0.0002	0.038, 0.013	0.636	0.302
4	0.921 ± 0.03	0.894 ± 0.05	0.061, 0.024	0.062, 0.024	0.977	0.984

Table 8.3: Area under curve (AUC) and partial AUC evaluated at 90% and 95% sensitivities along with the results of the statistical test between the partial AUC values obtained under the mono and stereo viewing modes ($p < 0.05$ indicates statistical significance).

Chapter 9

Conclusion and Future Work

In this dissertation, we have made an attempt to extend the barrier in computer-aided analysis and interpretation of breast imaging data. CADe systems for mammography are of great importance since they have been shown in several studies to positively affect the outcome of mammographic interpretation. Correct and timely mammographic interpretation could help save lives. One area where CADe systems for mammography need improvement is in the detection of spiculated lesions. Current CADe systems for mammography lack good algorithms for explicitly capturing and annotating spicule information, which is a non-trivial task since spicules are fine scale curvilinear structures that are often suppressed in the surrounding breast parenchyma. Spicule annotation can not only yield useful discriminative information about the suspect lesion location on the mammogram, but can also provide rich visual evidence to the interpreting radiologist to make the right follow-up decision. Our snakules algorithm for spicule annotation is motivated by these reasons and we have shown via observer studies with experienced radiologists that the algorithm is reliable in annotating obvious spicules.

In spite of mammography being the first choice modality for screening asymptomatic women for signs of early breast cancer, mammography is not perfect. One of the key limitations of mammography is the presence of anatomical noise caused by the superposition of overlapping, out-of-plane tissue structures, which hinders both humans and computer-based systems in making the right decisions. This has spurred the radiology community to investigate new 3-D breast imaging modalities for breast cancer screening and diagnosis. The last 15 years have seen the advent of breast tomosynthesis, stereo mammography, and breast CT, all of which are 3-D modalities that promise to improve the detection performance offered by mammography.

The advent of these new modalities has opened the door for the development of computational tools for interpreting the imaging data acquired from these modalities. In this dissertation, we have dealt with the problem of disparity estimation for stereo mammography. Disparity estimation is the first step towards building computational tools for interpreting stereo mammography data. Computational tools for stereo mammography could play a role in various interpretation tasks such as measuring lesion properties in 3-D, estimating the depth at which a tumor is located, providing highly localized 3-D information for biopsy, and quantifying parenchymal patterns and breast density in 3-D. However, the problem of disparity estimation for stereo images is generally a complex one made difficult by factors such as non-unique matches, half occlusions, and transparency. The problem is further compounded in the case of stereo mammography due to the presence of many singularities in the form of curvilinear structures of various lengths, widths, and tortuosities, which exhibit a complex occlusion pattern. By comparison, optical images of typical natural scenes are largely comprised of piece-wise smooth surfaces. Fine scale curvilinear structures such as spicules, blood vessels, and ducts are salient structures in the breast and the disparity along these structures has to be reliably estimated in order to correctly position them in 3-D. Motivated by this goal, we have developed a new stereo model that estimates a dense disparity map from a pair of stereo mammogram images. A key component of the proposed stereo model is a new singularity index that we have developed for reliably detecting curvilinear singularities in images. In-depth theoretical analysis addressing questions such as robustness to noise and extensive experiments with real images highlight the advantages of the new singularity index over some of the state of the art ridge strength measures proposed in the computer vision literature. The stereo model employs this singularity index to better estimate the disparity along curvilinear structures. Experiments on synthetic stereo pairs and on real stereo mammogram images suggest that the proposed stereo model is better than the canonical stereo model that relies on the piece-wise smooth assumption of the 3-D scene to constrain the resulting disparity.

The power of stereo visualization is not limited to stereo mammography alone. In fact, any 3-D x-ray based breast imaging modality in which multiple projection images are acquired at different angles that are closely spaced (say 5-10 degrees) can benefit from stereo visualization. Stereo visualization reveals the true 3-D structure of the breast. In this dissertation, we have demonstrated the use of stereo visualization for interpreting breast tomosynthesis projection images. We showed that despite the low radiation dose at which each projection image is acquired, 3-D perception is still enabled when the projection images are viewed stereoscopically. We also showed that stereo viewing of low dose tomosynthesis projection images could produce equivalent or better mass detection performance than monoscopic viewing of the projection images.

There are a number of research questions that have been set up by this dissertation that can be addressed going forward. For instance, one aspect of the snakules algorithm that needs improvement is the annotation of curvilinear structures that are non-spicules but are directed towards the central mass. A possible approach to address this problem could be to combine spicule identification with the snakules algorithm. For example, spicule identification could be performed as a post-processing step to filter out snakule annotations that are not spicules. For this, one could develop classification models based on features extracted from intensity cross-sectional profiles of curvilinear structures that could help differentiate between spicules and non-spicules. The work by [111] is relevant for this idea. Secondly, it would be interesting to evaluate snakules in conjunction with a good CADe algorithm for spiculated lesions on mammographic images. The motivation here would be to adopt snakules to improve the specificity of the CADe algorithm.

Similarly, our work on disparity estimation for stereo mammography has a number of interesting areas for future research. For instance, the proposed stereo model does not consider the vergent geometry of the stereo mammographic image acquisition system. Also, the proposed model does not handle occlusion explicitly. Explicitly integrating the epipolar constraint by taking into account the vergent geometry of the stereo system and handling occlusion within the optimization framework might yield better disparity estimates for stereo mammogram image pairs. Examples of relevant work for these areas include [99] and [8]. Also, better and faster optimization algorithms such as those based on graph cuts for optimizing the proposed stereo model can be explored. For example, there has been recent work on optimizing energy functionals containing higher order priors using graph cuts based techniques [106] and it would be interesting to explore these in the context of the proposed model. Ultimately, the purpose of disparity estimation algorithms for stereo mammogram image pairs is to recover depth estimates for the structures within the breast and use these depth estimates to quantify anatomically relevant measurements. This remains to be explored in the future.

Also, it has to be noted that the new singularity index that we have developed is a general purpose index and can be used for any application where simultaneous detection of curvilinear structures and suppression of edges is important. Here, we have used this index to detect curviinear structures in mammograms in the context of the disparity estimation problem. Other example applications for the index include segmentation of neuronal axons and filaments in biological specimens that have been imaged using techniques such as confocal microscopy, and the detection of roads and river deltas depicted on satellite images. Also, it would be interesting to study what types of higher order singularities can be detected within the framework of the generalized k^{th} order index.

Finally, our stereo observer study involving tomosynthesis projection images can be expanded by including reconstructed slices as a reading mode to assess how stereo viewing of projection images compares against the standard cine or slice-by-slice viewing of reconstructed slices. It would be interesting to measure both efficiency (e.g., in terms of time taken to interpret a case) and detection performance, especially of subtle breast lesions for a large number of experienced breast imaging radiologists.

Bibliography

[1] E. H. Adelson, C. H. Anderson, J. R. Bergen, P. J. Burt, and J. M. Ogden. Pyramid methods in image processing. *RCA Engineer*, 29(6):33–41, 1984.

[2] I. Andersson, D. M. Ikeda, S. Zackrisson, M. Ruschin, T. Svahn, P. Timberg, and A. Tingberg. Breast tomosynthesis and digital mammography: a comparison of breast cancer visibility and BIRADS classification in a population of cancers with subtle mammographic findings. *Eur J Radiol*, 18(12):2817–2825, 2008.

[3] S. M. Astley. Evaluation of computer-aided detection (CAD) prompting techniques for mammography. *Br J Radiol*, 78(Special Issue):S20–S25, 2005.

[4] J. A. Baker, E. L. Rosen, J. Y. Lo, E. I. Gimenez, R. Walsh, and M. S. Soo. Computer-aided detection (CAD) in screening mammography: sensitivity of commercial CAD systems for detecting architectural distortion. *AJR. AJR Am J Roentgenol.*, 181(4):1083–1088, 2003.

[5] L. Barker, H. Rolka, D. Rolka, and C. Brown. Equivalence testing for binomial random variables: Which test to use? *AM Stat*, 55(4):279–287, 2001.

[6] S. T. Barnard. Stochastic stereo matching over scale. *Int. J. Comput. Vision*, 3(1):17–32, 1989.

[7] L. W. Bassett, C. Bent, J. W. Sayre, R. Marzan, A. Verma, and C. Porter. Breast imaging training and attitudes: update survey of senior radiology residents. *AJR Am J Roentgenol*, 197(1):263–9, 2011.

[8] R. Ben-Ari and N. Sochen. Stereo matching with mumford-shah regularization and occlusion handling. *IEEE Trans. Pattern Anal. Mach. Intell.*, 32(11):2071–2084, 2010.

[9] M. O. Berger. Snake growing. In *European conference on Computer vision*, Antibes, France, 1990. Springer-Verlag New York, Inc.

[10] R. L. Birdwell. The preponderance of evidence supports computer-aided detection for screening mammography. *Radiology*, 253(1):9–16, 2009.

[11] R. L. Birdwell, D. M. Ikeda, K. F. OShaughnessy, and E. A. Sickles. Mammographic characteristics of 115 missed cancers later detected with screening mammography and the potential utility of computer-aided detection. *Radiology*, 219(1):192–202, 2001.

[12] J. M. Boone, A. L. Kwan, K. Yang, G. W. Burkett, K. K. Lindfors, and T. R. Nelson. Computed tomography for imaging the breast. *J Mammary Gland Biol Neoplasia.*, 11(2):103–111, 2006.

[13] J. M. Boone, T. R. Nelson, K. K. Lindfors, and J. A. Seibert. Dedicated breast CT: radiation dose and image quality evaluation. *Radiology*, 221(3):657–667, 2001.

[14] A. C. Bovik, P. Maragos, and T. F. Quatieri. AM-FM energy detection and separation in noise using multiband energy operators. *IEEE Trans. Signal Process.*, 41(12):3245–3265, 1993.

[15] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Anal. Mach. Intell.*, 23(11):1222–1239, 2001.

[16] J. E. Bresenham. Algorithm for computer control of a digital plotter. *IBM Systems Journal*, 4(1):25–30, 1965.

[17] M. Z. Brown, D. Burschka, and G. D. Hager. Advances in computational stereo. *IEEE Trans. Pattern Anal. Mach. Intell.*, 25(8):993–1008, 2003.

[18] A. E. Burgess, F. L. Jacobson, and P. F Judy. Human observer detection experiments with mammograms and power-law noise. *Med Phys.*, 28(4):419–437, 2002.

[19] L. W. Burhenne, S. A. Wood, C. J. D'Orsi, S. A. Feig, D. B. Kopans, K. F. O'Shaughnessy, E. A. Sickles, L. Tabar, C. J. Vyborny, and R. A. Castellino. Potential contribution of computer-aided detection to the sensitivity of screening mammography. *Radiology*, 215(2):554–62, 2000.

[20] J. Canny. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 8(6):679–698, 1986.

[21] R. A. Castellino. Computer aided detection (CAD): an overview. *Cancer imaging*, 5(1):17–19, 2005.

[22] A. Cedilnik, K. Kosmelj, and A. Blejec. Ratio of two random variables: A note on the existence of its moments. *Metodoloski zvezki*, 3(1):1–7, 2006.

[23] D. M. Chelberg, J. Hsu, C. F. Babbs, Z. Pizlo, and E. J. Delp. Digital stereomammography. In *Proceedings of the 2nd international workshop on Digital Mammography*, 1994.

[24] J. T. Dobbins and D. J. Godfrey. Digital x-ray tomosynthesis: current state of the art and clinical potential. *Phys. Med. Biol.*, 48(19):R65–106, 2003.

[25] C. J. D'Orsi, L. W. Bassett, W. A. Berg, and et al. *BI-RADS: Mammography.* Breast Imaging Reporting and Data System: ACR BI-RADS-breast imaging atlas. American College of Radiology, Reston (VA), 4th edition, 2003.

[26] G. Egnal and R. P. Wildes. Detecting binocular half-occlusions: empirical comparisons of five approaches. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(8):1127–1133, 2002.

[27] J. J. Fenton, L. Abraham, S. H. Taplin, B. M. Geller, P. A. Carney, C. D'Orsi, J. G. Elmore, and W. E. Barlow. Effectiveness of computer-aided detection in community mammography practice. *J Natl Cancer Inst*, 103(15):1152–1161, 2011.

[28] J. J. Fenton, S. H. Taplin, P. A. Carney, L. Abraham, E. A. Sickles, C. D'Orsi, E. A. Berns, G. Cutter, R. E. Hendrick, W. E. Barlow, and J. G. Elmore. Influence of computer-aided detection on performance of screening mammography. *N Engl J Med*, 356(14):1399–1409, 2007.

[29] W. T. Freeman and E. H. Adelson. The design and use of steerable filters. *IEEE Trans. Pattern Anal. Mach. Intell.*, 13(9):891–906, 1991.

[30] T. W. Freer and M. J. Ulissey. Screening mammography with computer-aided detection: prospective study of 12,860 patients in a community breast center. *Radiology.*, 220(3):781–786, 2001.

[31] D. J. Getty, C. J. D'Orsi, and R. M. Pickett. Stereoscopic digital mammography: improved accuracy of lesion detection in breast cancer screening. In *Proceedings of the 9th international workshop on Digital Mammography*, IWDM '08, pages 74–79. Springer-Verlag, 2008.

[32] F. J. Gilbert, S. M. Astley, M. G. C. Gillan, O. F. Agbaje, M. G. Wallis, J. James, C. R.M. Boggis, and S. W. Duffy. Single reading with computer-aided detection for screening mammography. *N Engl J Med*, 359(16):1675–1684, 2008.

[33] W. F. Good, G. S. Abrams, V. J. Catullo, D. M. Chough, M. A. Ganott, C. M. Hakim, and D. Gur. Digital breast tomosynthesis: A pilot observer study. *AJR Am J Roentgenol*, 190(4):865–869, 2008.

[34] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products*. Academic Press, San Diego, CA, fifth edition, 1994.

[35] M. Gromet. Comparison of computer-aided detection to double reading of screening mammograms: Review of 231,221 mammograms. *AJR Am J Roentgenol*, 190(4):854–859, 2008.

[36] D. Gur, G. S. Abrams, D. M. Chough, M. A. Ganott, C. M. Hakim, R. L. Perrin, G. Y. Rathfon, J. H. Sumkin, M. L. Zuley, and A. I. Bandos. Digital breast tomosynthesis: observer performance study. *AJR Am J Roentgenol*, 193(2):586–591, 2009.

[37] D. Gur, A. I. Bandos, J. L. King, A. H. Klym, C. S. Cohen, C. M. Hakim, L. A. Hardesty, M. A. Ganott, R. L. Perrin, W. R. Poller, R. Shah, J. H. Sumkin, L. P. Wallace, and H. E. Rockette. Binary and multi-category ratings in a laboratory observer performance study: a comparison. *Med Phys*, 35(10):4404–9, 2008.

[38] K. L. Gwet. Computing inter-rater reliability and its variance in the presence of high agreement. *British Journal of Mathematical and Statistical Psychology*, 61(1):29–48, 2008.

[39] M. Heath, K.W. Bowyer, and D. Kopans. *Current status of the Digital Database for Screening Mammography.* Kluwer Academic Publishers, Dordrecht, 1998.

[40] M. Heath, K.W. Bowyer, D. Kopans, R. Moore, and Jr. Kegelmeyer, P. The digital database for screening mammography. In M. J. Yaffe, editor, *5th International Workshop on Digital Mammography*, pages 212–218, Toronto, Canada, 2001. Med Phys Publishing.

[41] H. Hirschmuller and D. Scharstein. Evaluation of stereo matching costs on images with radiometric differences. *IEEE Trans. Pattern Anal. Mach. Intell.*, 31(9):1582–1599, September 2009.

[42] C. R. Huang, P. C. Chung, T. Y. Lee, S. C. Yang, and S. K. Lee. Reconstruction and rendering of microcalcifications from two mammogram views by modified projective grid space (mpgs). *Comput Med Imaging Graph.*, 30(2):123–33, 2006.

[43] L. Isserlis. On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. *Biometrika*, 12(1/2):134–139, 1918.

[44] J. F. Kaiser. Some useful properties of teager's energy operators. In *IEEE Intl. Conf. Acoust., Speech, and Signal Process.*, volume 3, pages 149–152. IEEE, 1993.

[45] A. Karellas and S. Vedantham. Breast cancer imaging: a perspective for the next decade. *Med Phys*, 35(11):4878–97, 2008.

[46] N. Karssemeijer, A. M. Bluekens, D. Beijerinck, J. J. Deurenberg, M. Beekman, R. Visser, R. Van Engen, A. Bartels-Kortland, and M. J. Broeders. Breast cancer screening results 5 years after introduction of digital mammography in a population-based screening program. *Radiology*, 253(2):353–358, 2009.

[47] N. Karssemeijer and G.M. te Brake. Detection of stellate distortions in mammograms. *IEEE Trans. Med. Imag.*, 15(5):611–619, 1996.

[48] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: active contour models. *Intl. J. Comput. Vision*, 1(4):321–331, 1987.

[49] N. Kim, A. C. Bovik, and S. J. Aggarwal. Shape description of biological objects via stereo light microscopy. *IEEE Trans. Syst., Man, Cybern.*, 20(2):475–489, 1990.

[50] N. H. Kim and A. C. Bovik. A contour-based stereo matching algorithm using disparity continuity. *Pattern Recogn*, 21(5):505–514, 1988.

[51] S. J. Kim, W. K. Moon, S. Y. Kim, J. M. Chang, S. M. Kim, and N. Cho. Comparison of two software versions of a commercially available computer-aided detection (CAD) system for detecting breast cancer. *Acta radiol*, 51(5):482–90, 2010.

[52] E. M. Knorr and R. T. Ng. Finding intensional knowledge of distance-based outliers. In *VLDB Conference*, pages 211–222, 1999.

[53] I. Kokkinos, P. Maragos, and A Yuille. Bottom-up and top-down object detection using primal sketch features and graphical models. In *IEEE Conf. Comput. Vision and Pattern Recog.*, pages 1893 – 1900. IEEE, 2006.

[54] T. M. Koller, G. Gerig, G. Szekely, and D. Dettwiler. Multiscale detection of curvilinear structures in 2-D and 3-D image data. In *IEEE Intl. Conf. Comput. Vision*, pages 864–869. IEEE, 1995.

[55] D. B. Kopans. Screening mammography. *Lancet*, 342(8870):550, 1993.

[56] H. Larochelle, D. Erhan, A. Courville, J. Bergstra, and Y. Bengio. An empirical evaluation of deep architectures on problems with many factors of variation. In *Proc. of the Twenty-fourth International Conference on Machine Learning (ICML'07)*, pages 473–480. ACM, 2007.

[57] Bing Li and Scott T. Acton. Active contour external force using vector field convolution for image segmentation. *IEEE Trans. Image Process.*, 16(8):2096–2106, 2007.

[58] L. Li, F. Mao, W. Qian, and L.P. Clarke. Wavelet transform for directional feature extraction in medical imaging. In *IEEE International Conference on Image Processing*, volume 3, pages 500–503, 1997.

[59] L. Liberman, A. F. Abramson, F. B. Squires, J. R. Glassman, E. A. Morris, and D. D. Dershaw. The breast imaging reporting and data system: positive predictive value of mammographic features and final assessment categories. *AJR. AJR Am J Roentgenol.*, 171(1):35–40, 1998.

[60] T. Lindeberg. Edge detection and ridge detection with automatic scale selection. *Intl. J. Comput. Vision*, 30(2):117–154, 1998.

[61] T. Lindeberg. Feature detection with automatic scale selection. *Intl. J. Comput. Vision*, 30(2):77–116, 1998.

[62] S. Liu, C.F. Babbs, and E.J. Delp. Multiresolution detection of spiculated lesions in digital mammograms. *IEEE Trans. Image Process.*, 10(6):874–884, 2001.

[63] Y. Liu, L. K. Cormack, and A. C. Bovik. Statistical modeling of 3-D natural scenes with application to bayesian stereopsis. *IEEE Trans. Image Process.*, 20(9):2515–2530, 2011.

[64] S. Mallat and W. L. Hwang. Singularity detection and processing with wavelets. *IEEE Trans. Inf. Theory*, 38(2):617–643, 1992.

[65] R. March. Computation of stereo disparity using regularization. *Pattern Recogn. Lett.*, 8(3):181–187, 1988.

[66] R. March. A regularization model for stereo vision with controlled continuity. *Pattern Recogn. Lett.*, 10(4):259–263, 1989.

[67] D. Marr and E Hildreth. Theory of edge detection. *Proc. R. Soc. Lond. B 29*, 207(1167):187–217, 1980.

[68] G. S. Muralidhar, A. C. Bovik, J. D. Giese, M. P. Sampat, G. J. Whitman, T. M. Haygood, T. W. Stephens, and M. K. Markey. Snakules: A model-based active contour algorithm for the annotation of spicules on mammography. *IEEE Trans. Med. Imag.*, 29(10):1768–80, 2010.

[69] G. S. Muralidhar, A. C. Bovik, and M. K. Markey. Snakules: snakes that seek spicules on mammography. In *17th IEEE International Conference on Image Processing (ICIP)*, pages 4373–4376, Hong Kong, 2010.

[70] G. S. Muralidhar, A. C. Bovik, and M. K. Markey. Disparity estimation on stereo mammograms. *IEEE Trans. Med. Imag.*, submitted, 2012.

[71] G. S. Muralidhar, A. C. Bovik, and M. K. Markey. A new singularity index. In *IEEE Intl. Conf. Image Proc.* IEEE, 2012.

[72] G. S. Muralidhar, A. C. Bovik, and M. K Markey. Noise analysis of a new singularity index. *IEEE Trans. Signal Process.*, In Preparation, 2012.

[73] G. S. Muralidhar, A. C. Bovik, and M. K Markey. A steerable, multiscale singularity index. *IEEE Signal Process. Lett.*, 20(1):7–10, 2013.

[74] G. S. Muralidhar, T. Ganapathi, A. C. Bovik, M. K. Markey, T. M. Haygood, T. W. Stephens, and G. J. Whitman. Stereoscopic vs. monoscopic detection of masses on breast tomosynthesis images. In *Proc. of*

SPIE Medical Imaging: Image Perception, Observer Performance, and Technology Asessment, volume 8318, pages 07–1–07–7, 2012.

[75] Gautam S. Muralidhar, Mia K. Markey, and Alan C. Bovik. Snakules for automatic classification of spiculated mass locations on mammography. In *IEEE Southwest Symposium on Image Analysis and Interpretation*, pages 197–200, Austin, TX, USA, 2010.

[76] L. T. Niklason, B. T. Christian, L. E. Niklason, D. B. Kopans, D. E. Castleberry, B. H. Opsahl-Ong, C. E. Landberg, P. J. Slanetz, A. A. Giardino, R. Moore, D. Albagli, M. C. DeJule, P. F. Fitzgerald, D. F. Fobare, B. W. Giambattista, R. F. Kwasnick, J. Liu, S. J. Lubowski, G. E. Possin, J. F. Richotte, C. Y. Wei, and R. F. Wirth. Digital tomosynthesis in breast imaging. *Radiology*, 205(2):399–406, 1997.

[77] R. M. Nishikawa. Current status and future directions of computer-aided diagnosis in mammography. *Computerized Medical Imaging and Graphics*, 31(4-5):224–235, 2007.

[78] R. M. Nishikawa and J. J. Fenton. Computer-aided detection should be used routinely to assist screening mammogram interpretation. *Med Phys*, 39(9):5305–5307, 2012.

[79] S. G. Orel, N. Kay, C. Reynolds, and D. C. Sullivan. BI-RADS categorization as a predictor of malignancy. *Radiology.*, 211(3):845–50, 1999.

[80] A. Papoulis and S. U. Pillai. *Probability, Random Variables and Stochastic Processes*. McGraw-Hill, New York, NY, fourth edition, 2002.

[81] S. P. Poplack, T. D. Tosteson, C. A. Kogel, and H. M. Nagy. Digital breast tomosynthesis: Initial experience in 98 women with abnormal digital screening mammography. *AJR Am J Roentgenol*, 189(3):616–623, 2007.

[82] J. L. Prince and J. M. Links. Projection radiography. In *Medical imaging signals and systems*, pages 135–180. Pearson Prentice Hall, Upper Saddle River, 2006.

[83] S. Psarakis and J. Panaretos. On some bivariate extensions of the folded normal and the folded t distributions. *J. Appl. Statist. Sci.*, 10(2):119–136, 2001.

[84] W. Qian, L. Li, L. Clarke, R. A. Clark, and J. Thomas. Digital mammography: comparison of adaptive and nonadaptive CAD methods for mass detection. *Acad Radiol*, 6(8):471–480, 1999.

[85] E. Rafferty and L. T. Niklason. Comparison of FFDM with breast tomosynthesis to FFDM alone: performance in fatty and dense breasts. In *Tomosynthesis Imaging Symposium 2009: Frontiers in Research and Clinical Applications*, Duke University, Durham, NC, USA, 2009.

[86] E. A. Rafferty, L. Niklason, and E. Halpern. Assessing radiologist performance using combined full field digital mammography and breast to-

mosynthesis versus full-field digital mammography alone: results of a multi-center, multi-reader trial. In *RSNA*, 2007.

[87] R. M. Rangayyan and F. J. Ayres. Gabor filters and phase portraits for the detection of architectural distortion in mammograms. *Med Bio Eng Comput.*, 44:883894, 2006.

[88] W. Richards. Stereopsis and stereoblindness. *Experimental Brain Research*, 10(4):380–388, 1970.

[89] L. G. Roberts. Machine perception of three-dimensional solids. In *Optical and Electro-Optical Information Processing*, pages 159–197. MIT Press, 1965.

[90] X. Robin, N. Turck, A. Hainard, N. Tiberti, F. Lisacek, J.C. Sanchez, and M. Muller. proc: an open-source package for r and s+ to analyze and compare roc curves. *BMC Bioinformatics*, 12:77, 2011.

[91] B. Rosner. *Fundamentals of Biostatistics.* Duxbury Press, Belmont, CA, 6th edition, 2005.

[92] A. Sadaf, P. Crystal, A. Scaranelo, and T. Helbich. Performance of computer-aided detection applied to full-field digital mammography in detection of breast cancers. *Eur J Radiol*, 77(3):457–61, 2011.

[93] M. P. Sampat, A.C. Bovik, G.J. Whitman, and M.K. Markey. A model-based framework for the detection of spiculated masses on mammography. *Med Phys*, 35(5):2110–2123, 2008.

[94] M. P. Sampat, M. K. Markey, and A. C. Bovik. Computer-aided detection and diagnosis in mammography. In A.C. Bovik, editor, *Handbook of Image and Video Processing*, pages 1195–1217. Academic Press, 2nd edition, 2005.

[95] M. P. Sampat, G. J. Whitman, T. W. Stephens, L. D. Broemeling, N. A. Heger, A. C. Bovik, and M. K. Markey. The reliability of measuring physical characteristics of spiculated masses on mammography. *BR J Radiol*, 79(Special Issue 2):S134–S140, 2006.

[96] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *Int. J. Comput. Vision*, 47(1-3):7–42, 2002.

[97] P. Skaane. Studies comparing screen-film mammography and full-field digital mammography in breast cancer screening: updated review. *Acta Radiol.*, 50(1):3–14, 2009.

[98] P. Skaane, A. Kshirsagar, S. Stapleton, K. Young, and R. A. Castellino. Effect of computer-aided detection on independent double reading of paired screen-film and full-field digital screening mammograms. *AJR Am J Roentgenol*, 188(2):377–384, 2007.

[99] N. Slesareva, A. Bruhn, and J. Weickert. Optic flow goes stereo: A variational method for estimating discontinuity-preserving dense disparity maps. *Lect Notes Comput Sc*, 3663:33–40, 2005.

[100] C. Steger. An unbiased detector of curvilinear structures. *IEEE Trans. Pattern Anal. Mach. Intell.*, 20(2):113–125, 1998.

[101] J. Sun, N. N. Zheng, and H. Y. Shum. Stereo matching using belief propagation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 25(7):787–800, 2003.

[102] A. Tiedeu, C. Daul, P. Graebling, and D. Wolf. Correspondences between microcalcification projections on two mammographic views acquired with digital systems. *Comput Med Imaging Graph.*, 29(7):543–53, 2005.

[103] L. J. Webb, E. Samei, J. Y. Lo, J. A. Baker, S. V. Ghate, C. Kim, M. S. Soo, and R. Walsh. Comparative performance of multiview stereoscopic and mammographic display modalities for breast lesion detection. *Med. Phys.*, 38(4):1972–80, 2011.

[104] D. J. Williams and M. Shah. A fast algorithm for active contours and curvature estimation. *CVGIP: Image Understanding*, 55(1):14–26, 1991.

[105] P. Wing and M. H. Langelier. Workforce shortages in breast imaging: impact on mammography utilization. *AJR Am J Roentgenol*, 192(2):370–8, 2009.

[106] O. Woodford, P. Torr, I. Reid, and A. Fitzgibbon. Global stereo reconstruction under second-order smoothness priors. *IEEE Trans. Pattern Anal. Mach. Intell.*, 31(12):2115–2128, 2009.

[107] C. Xu and J. L. Prince. Snakes, shapes, and gradient vector flow. *IEEE Trans. Image Process.*, 7(3):359–369, 1998.

[108] S. K. Yang, W. K. Moon, N. Cho, J. S. Park, J. H. Cha, S. M. Kim, S. J. Kim, and J. G.i Im. Screening mammography-detected cancers: sensitivity of a computer-aided detection system applied to full-field digital mammograms. *Radiology*, 244(1):104–111, 2007.

[109] S. C. Yen and L. H. Finkel. Extraction of perceptually salient contours by striate cortical networks. *Vision Res*, 38(5):719–741, 1998.

[110] Y. Zheng, M. Wu, E. Cole, and E. D. Pisano. Online annotation tool for digital mammography. *Acad Radiol*, 11(5):566–572, 2004.

[111] R. Zwiggelaar, S. M. Astley, C. R. M. Boggis, and C. J. Taylor. Linear structures in mammographic images: detection and classification. *IEEE Trans. Med. Imag.*, 23(9):1077–1086, 2004.

[112] R. Zwiggelaar, T. C. Parr, J. E. Schumm, I. W. Hutt, C. J. Taylor, S. M. Astley, and C. R. Boggis. Model-based detection of spiculated lesions in mammograms. *Med Image Anal*, 3(1):39–62, 1999.

Vita

Gautam Sakleshpur Muralidhar was born in Mumbai, India on 24 August 1980. He received the Bachelor of Engineering degree in Electronics and Communications Engineering from the Visvesvaraya Technological University in 2002. He began his professional career as a software engineer in July 2002 at Oracle India Development Center in Hyderabad, India. He left Oracle in July 2003 to begin work as a software engineer at Philips Medical Systems, a division of Philips Electronics India Ltd., in Bangalore, India. Gautam then joined The University of Texas at Austin in 2007 to begin his Ph.D. studies in Biomedical Engineering under the supervision of Dr. Alan C. Bovik and Dr. Mia K. Markey, where he also received the M. S. degree in Biomedical Engineering in May 2009. Gautam was awarded The University of Texas Graduate School Continuing Fellowship for the academic year 2011-2012. He is a student member of the Institute of Electrical and Electronics Engineers (IEEE) and International Society for Optics and Photonics (SPIE).

Permanent address: 510, "Gokul", 4th Cross Road
Banashankari 3rd Stage, 3rd Phase, 3rd Block
Bangalore - 560085, India.

This dissertation was typeset with LaTeX[†] by the author.

[†] LaTeX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's TeX Program.